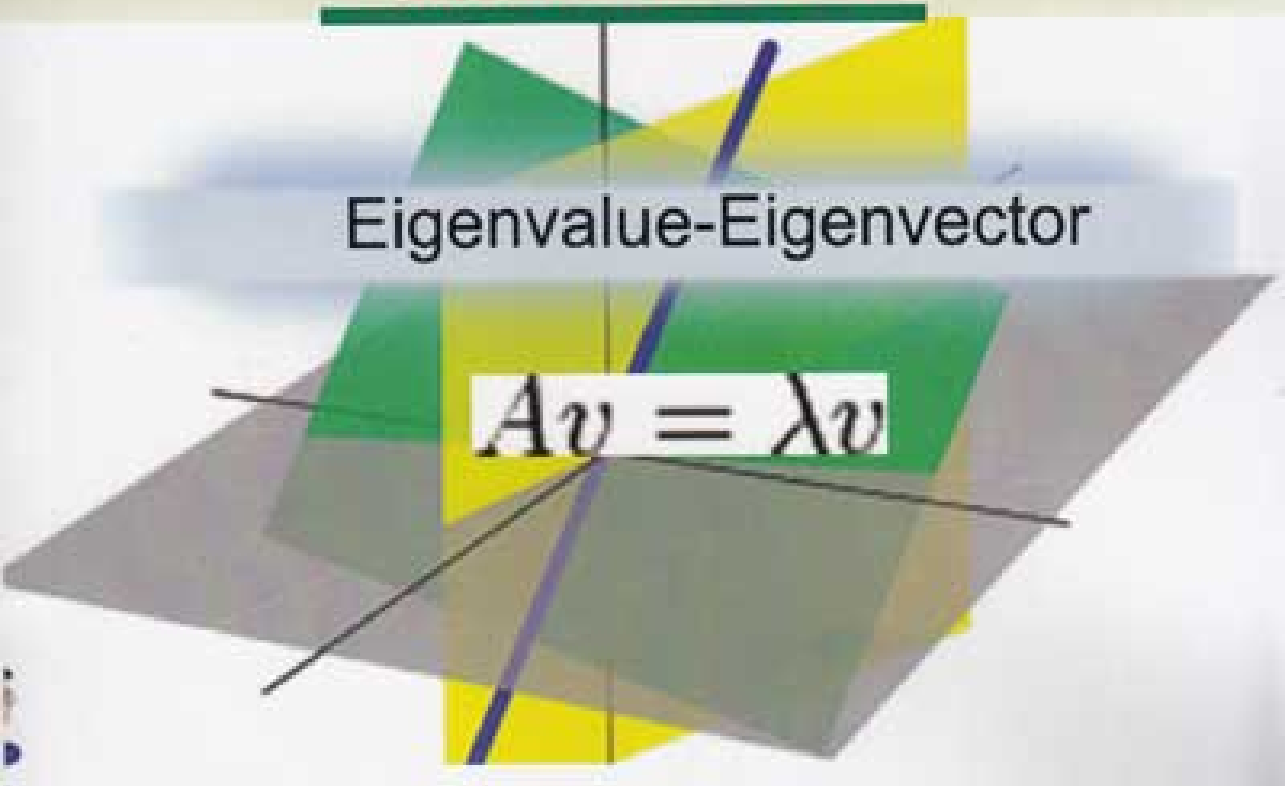
WITH MATLAB

LINEAR ALGEBRA

RANGRAO S. BHAMARE
PRAVIN S. WALDHE
ANAGHA R. MEDHEKAR



NIRALI PRAKASHAN

A Book Of

With Matlab

LINEAR ALGEBRA

For
M.Sc. Statistics and S.Y.B.Sc. / T.Y.B.Sc. / M.Sc. Mathematics

(Undergraduate and Post Graduate Courses in Mathematics & Statistics)

Prof. Rangrao S. Bhamare
M.Sc. M.Phil
Ex-Head, P.G. Dept. of Mathematics
New Arts, Science and Commerce College
Ahmednagar.

Prof. Pravin S. Waldhe
M.Sc., SET
Assistant Professor
Department of Statistics
Modern College of Arts, Science and Commerce
Pune - 411005.

Prof. Anagha R. Medhekar
M.Sc. (Statistics), B.Ed. (Mathematics and Science)

Price ₹ 400.00

NIRALI PRAKASHAN
ADVANCEMENT OF KNOWLEDGE

N0840

Linear Algebra **ISBN : 978-93-86943-17-0**

First Edition : August 2017

© : Author

The text of this publication, or any part thereof, should not be reproduced or transmitted in any form or stored in any computer storage system or device for distribution including photocopy, recording, taping or information retrieval system or reproduced on any disc, tape, perforated media or other information storage device etc., without the written permission of Author with whom the rights are reserved. Breach of this condition is liable for legal action.

Every effort has been made to avoid errors or omissions in this publication. In spite of this, errors may have crept in. Any mistake, error or discrepancy so noted and shall be brought to our notice shall be taken care of in the next edition. It is notified that neither the publisher nor the author or seller shall be responsible for any damage or loss of action to any one, of any kind, in any manner, therefrom.

Published By :

NIRALI PRAKASHAN

Abhyudaya Pragati, 1312, Shivaji Nagar,
Off J.M. Road, PUNE – 411005
Tel - (020) 25512336/37/39, Fax - (020) 25511379
Email : niralipune@pragationline.com

Printed By :

STAR COPIERS PVT. LTD.

Kumthekar Road, Sadashiv Peth,
PUNE - 411 030
Tel - (020) 24479201

☞ DISTRIBUTION CENTRES

PUNE

Nirali Prakashan : 119, Budhwar Peth, Jogeshwari Mandir Lane, Pune 411002, Maharashtra
Tel : (020) 2445 2044, 66022708, Fax : (020) 2445 1538
Email : bookorder@pragationline.com, niralilocal@pragationline.com

Nirali Prakashan : S. No. 28/27, Dhyari, Near Pari Company, Pune 411041
Tel : (020) 24690204 Fax : (020) 24690316
Email : dhyari@pragationline.com, bookorder@pragationline.com

MUMBAI

Nirali Prakashan : 385, S.V.P. Road, Rasdhara Co-op. Hsg. Society Ltd.,
Girgaum, Mumbai 400004, Maharashtra
Tel : (022) 2385 6339 / 2386 9976, Fax : (022) 2386 9976
Email : niralimumbai@pragationline.com

☞ DISTRIBUTION BRANCHES

JALGAON

Nirali Prakashan : 34, V. V. Golani Market, Navi Peth, Jalgaon 425001,
Maharashtra, Tel : (0257) 222 0395, Mob : 94234 91860

KOLHAPUR

Nirali Prakashan : New Mahadvar Road, Kedar Plaza, 1st Floor Opp. IDBI Bank
Kolhapur 416 012, Maharashtra. Mob : 9850046155

NAGPUR

Pratibha Book Distributors : Above Maratha Mandir, Shop No. 3, First Floor,
Rani Jhanshi Square, Sitabuldi, Nagpur 440012, Maharashtra
Tel : (0712) 254 7129

DELHI

Nirali Prakashan : 4593/15, Basement, Aggarwal Lane 15, Ansari Road, Daryaganj
Near Times of India Building, New Delhi 110002
Mob : 09555778814

BENGALURU

Pragati Book House : House No. 1, Sanjeevappa Lane, Avenue Road Cross,
Opp. Rice Church, Bengaluru – 560002.
Tel : (080) 64513344, 64513355,Mob : 9880582331, 9845021552
Email:bharatsavla@yahoo.com

CHENNAI

Pragati Books : 9/1, Montieth Road, Behind Taas Mahal, Egmore,
Chennai 600008 Tamil Nadu, Tel : (044) 6518 3535,
Mob : 94440 01782 / 98450 21552 / 98805 82331,
Email : bharatsavla@yahoo.com

Note: Every possible effort has been made to avoid errors or omissions in this book. In spite this, errors may have crept in. Any type of error or mistake so noted, and shall be brought to our notice, shall be taken care of in the next edition. It is notified that neither the publisher, nor the author or book seller shall be responsible for any damage or loss of action to any one of any kind, in any manner, therefrom. The reader must cross check all the facts and contents with original Government notification or publications.

Dedicated to

Our Dear Parents

Authors

Preface ...

We have great pleasure to present this book, "Linear Algebra" for the students of M.Sc. Statistics, However, this book can be used as reference book for "Linear Algebra" courses taught at undergraduate and postgraduate in Mathematics and Statistics.

The content of the book mostly covers the syllabi of P.G. courses for "Linear Algebra" of all universities in India. A part of the book : chapters 1, 3, 4, 5, 6 can be taken as 'Linear Algebra' course at U.G. in any University in India.

We are very much encouraged by the overwhelming response for the text books written by us in the last several years. We tried to take utmost care of the need of the students as well as teachers. We have taken care to present the matter systematically. The book contains several solved examples for better understanding of the students. An ample number of graded problems are included in the exercises.

The inclusion of Matlab Commands is an additional feature of this book. It will help the students to compute the problems in Linear Algebra on the topics such as : g-inverse, Eigen values, Eigen vector and Quadratic Forms etc.

We are thankful to Shri. Dineshbhai Furia, Shri. Jignesh Furia, Mr. Santosh Bare, Mrs. Manasi Pingle and the staff of the Nirali Prakashan, for the great efforts that they have taken to publish the book in time. Needless to say the great encouragement and support by our family members and well-wishers provided us the source of energy to complete this book.

We welcome the valuable suggestions from our colleagues and readers for the improvement of the book.

Authors

Syllabus ...

UNIT - I

Solution of a system of homogeneous and non-homogeneous linear equations. Gauss seidal and Gauss elimination method.

UNIT - II

Vector space, subspace, linear dependence and independence, basis of vector space, dimension of a vector space, orthogonal and orthonormal vectors, orthonormal basis, Gram-Schmidt orthogonalization, Matrix algebra, special types of matrices, orthogonal matrix, idempotent matrix, partitioned matrices, elementary operations, rank of a matrix, inverse of a matrix.

UNIT - III

g-inverse, Moore-Penrose g-inverse.

UNIT - IV

Characteristic roots of a matrix, right and left characteristic vectors, properties of characteristic roots and vectors, algebraic and geometric multiplicities, spectral decomposition, n^{th} power of a matrix, Cayley-Hamilton theorem.

UNIT - V

Quadratic forms, definition, reduction and classification, simultaneous reduction of two quadratic forms, maxima and minima of ratio of quadratic forms.

•••

Contents ...

1. Linear Equations and Matrices **1.1 – 1.76**

1.1 Systems of Linear Equations 1.1

1.1.1 System of Linear Equations 1.3

1.1.2 Solution of a System of Linear Equations 1.3

1.1.3 Homogeneous and Non-Homogeneous System of Linear Equations 1.4

1.1.4 Definition 1.7

1.2 Matrices 1.8

1.2.1 Definition 1.8

1.2.2 Types of Matrices 1.8

1.3 Operations on Matrices 1.12

1.4 Block Multiplication of Matrices 1.22

1.5 Elementary Operations 1.33

1.6 Applications of Row-reduction for Solving Systems of Linear Equations 1.35

1.7 Elementary Matrices 1.49

1.8 LU-Factorization 1.68

2. Generalised Inverse **2.1 – 2.30**

2.1 Introduction 2.1

2.2 Generalized Inverse : Definition 2.2

2.3 Generalized inverse of Arbitrary matrices 2.5

2.4 Applications of generalized inverse in statistics 2.9

2.5 Moore-penrose generalized inverse 2.11

2.6 Existence of generalized inverse by Different Methods 2.13

2.7 Singular Value Decomposition (SVD) 2.18

2.8 SVD Algorithm 2.19

2.9 Full Rank Factorization Method 2.21

3. Vector Spaces **3.1 – 3.86**

3.1 Introduction of Real Vector Space 3.2

3.2 Subspaces 3.18

3.3 Linear Combination and Spanning Sets 3.26

3.4 Linear Dependence and Independence 3.36

3.5 Basis and Dimensions 3.47

3.6 Row Space, Column Space, Null Space, Rank and Nullity 3.64

4. Eigenvalues and Eigenvectors **4.1 – 4.54**

4.1 Introduction 4.2

4.2 Characteristic Polynomials of Degree 2 and 3 4.3

4.3 Cayley Hamilton Theorem 4.16

4.4 The Construction of Orthogonal Matrices 4.26

4.5 Diagonalization of Matrices 4.28

4.6 Minimal Polynomial and Minimal Equation of a Matrix 4.40

4.7 Derogatory and Non-derogatory Matrices 4.43

5. Linear Transformations **5.1 – 5.64**

5.1 Introduction 5.1

5.2 The Definition of Linear Transformation and Examples 5.2

5.3 Sum, Scalar Multiple and Composition of Linear Transformation 5.10

5.4 Kernel and Range 5.14

5.5 Inverse of a Linear Transformation 5.33

5.6 Matrix of a Linear Transformation 5.42

6. Inner Product Spaces **6.1 – 6.64**

6.1 Introduction 6.2

6.2 Dot Product of Euclidean Inner Product in R^n 6.2

6.3 General Inner Product 6.7

6.4 Angle and Orthogonality in an Inner Product Space 6.25

6.5 Orthonormal Bases, Gram-Schmidt Process 6.34

6.6 Rank and Nullity : Sylvester Inequality 6.52

7. Quadratic Forms **7.1 – 7.32**

7.1 Introduction of Quadratic Form 7.2

7.2 Singular and Non-singular Linear Transformations 7.10

7.3 Linear Transformations of Quadratic Form on Field F 7.12

7.4 Congruence of Quadratic Forms and Matrices 7.12

7.5 Elementary Congruent Transformations of Non-singular Matrix 7.13

7.6 Elementary Congruent Transformations 7.14

7.7 Congruent Reduction of a Symmetric Matrices 7.14

7.8 Congruence Reduction of Skew-symmetric Matrices 7.18

7.9 Properties of Definite, Semi-definite and Indefinite Forms 7.20

Appendix **A.1 – A.67**

Model Question Papers **M.1 – M.17**

•••

Chapter **1**...

Linear Equations and Matrices

Synopsis

1.1 Systems of Linear Equations
  1.1.1 System of Linear Equations
  1.1.2 Solution of a System of Linear Equations
  1.1.3 Homogeneous and Non-Homogeneous System of Linear Equations
  1.1.4 Definition
1.2 Matrices
  1.2.1 Definition
  1.2.2 Types of Matrices
1.3 Operations on Matrices
1.4 Block Multiplication of Matrices
1.5 Elementary Operations
1.6 Applications of Row-reduction for Solving Systems of Linear Equations
1.7 Elementary Matrices
1.8 LU-Factorization

Objectives

At the end of this chapter you will be able to understand :

(i) Concept of system of linear equations.

(ii) Echelon form of a matrix.

Carl Friedrich Gauss

Johann Carl Friedrich Gauss (/gaʊs/; German: Gauß, pronounced Latin: Carolus Fridericus Gauss) (30 April 1777 – 23 February 1855) was a German mathematician who contributed significantly to many fields, including number theory, algebra, statistics, analysis, differential geometry, geodesy, geophysics, mechanics, electrostatics, astronomy, matrix theory, and optics.

Sometimes referred to as the Princeps mathematicorum Latin, "the foremost of mathematicians") and "greatest mathematician since antiquity", Gauss had an exceptional influence in many fields of mathematics and science and is ranked as one of history's most influential mathematicians.

1.1 Systems of Linear Equations

The system of linear equations and their solution form the major topics in linear algebra.

We know that a line in a xy-plane can be represented algebraically by the equation of the form

$$ax + by = c,$$

where a, b and c are constants and x and y are variables (unknowns). This type of equation is called a linear equation in the variables x and y. More generally, we define a *linear equation* in the n-variables, $x_1, x_2, ..., x_n$ to be the expression of the form

$$a_1x_1 + a_2x_2 + ... + a_n x_n = b$$

where, $a_1, a_2, ..., a_n$ and b are real constants. The variables are also called the *unknowns*.

Example 1 : The following are linear equations

$3x - y = -7$ $\qquad$ $2x_1 + 3x_2 - 4x_3 + 7x_4 = -1$

$x = \frac{1}{3}\ y - z + 5$ $\qquad$ $x_1 + 2x_2 + ... + n\ x_n = 1$

Note : The linear equations does not involve any product or roots of variables. The following are *not* linear equations :

$x^2 + 2x + 3y = 5$ $\qquad$ $xy - y + z + 5 = 0$

$x - \cos y = 2$ $\qquad$ $\sqrt{x} + y + \sqrt{2} = 3$

but the equation $\sqrt{2}\ x - y + \sqrt[3]{3}\ z = \sqrt{5}$ is a linear equation because all the variables are linear form.

A Solution of a linear equation : If $a_1x_1 + a_2x_2 + ... + a_nx_n = b$ … (i)

is a linear equation, then the sequence of numbers $t_1, t_2, ..., t_n$, which satisfies equation (i), that is, when $x_1, x_2, ..., x_n$ are replaced by $t_1, t_2, ..., t_n$, in equation (1), the equation is satisfied, is called the *solution* of this linear equation. The set of all solutions of the equation is called a *solution set* or sometimes the *general solution.*

Example 2 : Find the solution set of

(a) $x + y = 5$

(b) $x_1 + 5x_2 - 3x_3 = -1$

Solution : (a) We observe that x = 2, y = 3 and x = – 1, y = 6 are clearly two solutions of the equation. To find set of all solutions, we can assign an arbitrary value to y and solve for x. So let us assign y an arbitrary value t, we obtain x = 5 – t, y = t.

These two formulae for x and y describe the solution set in terms of the arbitrary parameter t. The particular solution x = 2 and y = 3 can be obtained by substituting specific value for t, say t = 3. And the second solution is obtained by taking t = 6.

Note that we can assign arbitrary value to x and then obtain the value of y in terms of the parameter.

(b) To find solution set in this case we can assign arbitrary values to any two variables and solve for the third variable. In particular, if we assign arbitrary values s and t to x_2 and x_3, respectively, we obtain x_1 as

$$x_1 = -1 - 5s + 3t$$

Thus, the solution set of the equation (b) is

$$x_1 = -1 - 5s + 3t,\ x_2 = s,\ x_3 = t, \text{ with } t \in R.$$

Note that $x_1 = 2$, $x_2 = 0$ and $x_3 = 1$ is a particular solution obtained by taking s = 0 and t = 1.

1.1.1 System of Linear Equations

A finite set of linear equations in the n variables $x_1, x_2, ..., x_n$ is called a *system of linear equations*. A system of m linear equations in n variables $x_1, x_2, ..., x_n$ is usually written as

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ . \quad . \quad ..... \quad . \quad . \\ . \quad . \quad ..... \quad . \quad . \\ . \quad . \quad ..... \quad . \quad . \\ a_{m1}x_1 + a_{m2}x_2 + ..., + a_{mn}x_n &= b_m \end{aligned}\right\} \quad \ldots (1)$$

where a_{ij}, i = 1, 2, ..., m, j = 1, 2, ..., n and $b_1, b_2, ..., b_m$ are real numbers.

Example 3 :

$$\begin{aligned} x_1 - x_2 + 2x_3 - x_4 &= -1 \\ 2x_1 + x_2 - 2x_3 - 2x_4 &= -2 \\ -x_1 + 2x_2 - 4x_3 + x_4 &= 1 \end{aligned}$$

is a system of three linear equations in *four* variables.

Example 4 :

$$\begin{aligned} x - y &= 4 \\ -3x + 3y &= 8 \end{aligned}$$

is a system of *two* equations in *two* variables.

Example 5 : 2x + 3y = 4 is a system of *single* equation in *two* variables.

1.1.2 Solution of a System of Linear Equations

A sequence of n real numbers $t_1, t_2, ..., t_n$ is called a *solution* of the given system of m equations in n-variables (as in (1)). If $x_1, x_2, ..., x_n$ are replaced by $t_1, t_2, ..., t_n$ respectively, all the equations in the system are satisfied.

Observe that $x_1 = 0$, $x_2 = 2$, $x_3 = 1$ and $x_4 = 1$ is a solution of the system of linear equations in example (3) as these values satisfy all the three equations. In

case of example (4), there is no solution, because the second equation can be written as $x - y = -\frac{8}{3}$. Hence the resulting equivalent system

$$\begin{aligned} x - y &= 4 \\ x - y &= \frac{-8}{3} \end{aligned}$$

has contradictory equations. In example (5) it is evident that the system has solution.

$x = t$, $y = \frac{4}{3} - \frac{2}{3}\ t$ or $x = 2 - \frac{3}{2}\ t$, $y = t$ for $t \in \mathbb{R}$. Hence for the system in example (5), there are infinitely many solutions.

Example 6 : The system

$$\begin{aligned} x - 2y &= 1 \\ 2x - 3y &= -1 \end{aligned}$$

has only one solution namely $x = -5$ and $y_2 = -3$. Note that these two equations represent two straight lines in plane, which intersect in only one point, whose co-ordinates satisfy both the equations simultaneously.

Remark : From above examples, we observe that every system of linear equations has either *no solutions, exactly one solution* or *infinitely many solutions.*

Let us define two types of systems of linear equations.

1.1.3 Homogeneous and Non-homogeneous System of Linear Equations

A system of linear equations in which all the right hand constants are zero is called a *homogeneous* system of linear equations. If *not* all right hand side constants are zero, the system is called a *non-homogeneous* system of linear equations.

Example 7 :

$$\begin{aligned} x_1 + 3x_2 - x_3 &= 0 \\ x_2 - 4x_3 &= 0 \\ 5x_3 &= 0 \end{aligned}$$

is a homogeneous system of linear equations. Clearly $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ is solution of this system. In fact, this is the only solution.

Example 8 :

$$\begin{aligned} 2x_1 - 3x_2 &= 0 \\ 4x_1 - 6x_2 &= 0 \end{aligned}$$

is a homogeneous system of linear equations. Observe that the system is equivalent to a single equation $2x_1 - 3x_2 = 0$, since second equation is obtained by multiplying first equation by 2. In this example, we can assign arbitrary value to x_2 say t, then

$x_1 = \frac{3}{2}\ t$, $x_2 = t$, $t \in R$ is a general solution of this system.

$x_1 = \frac{3}{2}$, $x_2 = 1$, $x_1 = 3$, $x_2 = 2$ are particular solutions of this system as these values satisfy both the equations of the system. It is obvious that there are *infinitely many solutions* for this system.

Example 9 :
$$\begin{aligned} x_1 + x_2 + x_4 &= 0 \\ x_2 + x_3 &= 0 \\ x_4 &= -1 \end{aligned}$$

is a *non-homogeneous* system of linear equations, since all right hand side constants are not zero. This system can be written as

$$\begin{aligned} x_1 &= x_3 - x_4 \qquad (\because x_2 = -x_3) \\ x_2 &= -x_3 \\ x_4 &= -1 \end{aligned}$$

We can assign arbitrary value t for x_3, and we obtain solution of the system as

$$\begin{aligned} x_1 &= t + 1 \qquad (\because x_4 = -1) \\ x_2 &= -t \\ x_3 &= t \\ x_4 &= -1, \qquad t \in \mathbb{R} \end{aligned}$$

This is the general solution of the system. Note that $x_1 = 2$, $x_2 = -1$, $x_3 = 1$, $x_4 = -1$ is a particular solution for the system obtained by putting $t = 1$.

Note : The homogeneous system has always atleast one solution : $x_1 = 0$, $x_2 = 0$, ..., $x_n = 0$, which is called *trivial* or *zero solution.* Thus, in case of homogeneous system of linear equations, the system has either *infinitely many solutions* or has only *trivial solution.*

Matrix representation of systems of linear equations :

An arbitrary *non-homogeneous system* of m linear equations in n unknowns is written as

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n &= b_2 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n &= b_m \end{aligned}\right\} \quad \ldots (1)$$

where *not all* b_i's are zeros.

Similarly, an arbitrary *homogeneous system* of m linear equations in n unknowns is written as

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n &= 0 \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n &= 0 \end{aligned}\right\} \quad \ldots (2)$$

Now, using equality of matrices, the system (1) can be written as

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

where, in left hand side column matrix, the expressions are single entries. Further, using multiplication of matrices we write above equality as

$$\begin{bmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Further this can be written in precise as

$$AX = B \qquad \ldots (3)$$

where $A = \begin{bmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{bmatrix}$

which is called the *matrix of coefficients* of the system. And

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

are column matrices of *unknowns* and *right hand side constants* of the system respectively. Note that X is of order $n \times 1$ and B is of order $m \times 1$.

The matrix equation (3) is called the *matrix representation of the non-homogeneous system of linear equations.*

Similarly, the homogeneous system of linear equations given in (2) can be written in matrix form as : $AX = 0$ … (4)

where, A is the coefficient matrix and X is a column matrix of unknowns, O is the column matrix of zeros of order $m \times 1$.

Definition : Given a non-homogenous system of linear equations (1), the block matrix

$$[A \mid B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & ..... & a_{1n} & b_1 \\ a_{21} & a_{22} & ..... & a_{2n} & b_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{m1} & a_{m2} & ..... & a_{mn} & b_m \end{array}\right]$$

is called the *augmented matrix* of the given system.

Example 3 : Consider the following system of linear equations

$$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{aligned}$$

This system can be written in matrix form as

$$AX = B$$

where $A = \begin{bmatrix} 2 & 2 & 2 \\ -2 & 5 & 2 \\ 8 & 1 & 4 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $B = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

The augmented matrix for this system is

$$[A \mid B] = \left[\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ -2 & 5 & 2 & 1 \\ 8 & 1 & 4 & -1 \end{array}\right]$$

1.1.4 Definition

A system of linear equations is called *consistent* if it has atleast one solution, otherwise it is called *inconsistent.*

Note : From above note, it is clear that a *homogeneous* system of linear equations is always *consistent*.

We will see more details about solutions (Gaussian elimination method) of the system of linear equations in later section.

1.2 Matrices

1.2.1 Definition

A rectangular array of m × n real numbers with m rows and n columns enclosed by square brackets (or by parentheses) is called a *matrix of order m × n* or m × n matrix. The numbers in the array are called the entries in the matrix.

Example 10 : The following are matrices.

$$\begin{bmatrix} 2 & \frac{1}{2} & \frac{1}{3} \\ 0 & \pi & 5 \\ e & 3 & -1 \end{bmatrix} \quad \begin{bmatrix} 3 & 0 \\ -1 & 4 \\ 1 & 2 \end{bmatrix} \quad \begin{bmatrix} 4 & 1 & 2 & 0 \end{bmatrix} \quad \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

The first is the matrix of order 3 × 3, the second is the matrix of order 3 × 2, third is the matrix of order 1 × 4 and the fourth is the matrix of order 2 × 1.

Notation : Usually capital letter denote matrices and lowercase letters denote numerical entries of matrix. Thus,

$$A = \begin{bmatrix} x & y & z \\ u & v & w \end{bmatrix}$$

is a matrix A of order 2 × 3 with entries x, y, z, u, v and w as real numbers.

If A is a matrix, we will use a_{ij} to denote the entry that occurs in i^{th} row and j^{th} column of A. Thus, a general 4 × 3 matrix is written as

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

Usually, we match the letters used to denote a matrix and the letter used to denote its entries. Thus, for a matrix X we use x_{ij} for the entry in i^{th} row and j^{th} column. The general m × n matrix is written as :

$$X = \begin{bmatrix} x_{11} & x_{12} & \ldots\ldots & x_{1n} \\ x_{21} & x_{22} & \ldots\ldots & x_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{m1} & x_{m2} & \ldots\ldots & x_{mn} \end{bmatrix} \quad \text{or } X = [x_{ij}]_{m \times n}$$

1.2.2 Types of Matrices

Square matrix : A matrix A = $[a_{ij}]$ with n rows and n-columns (or n × n matrix) is called *square matrix of order n* and the entries a_{11}, a_{22}, ..., a_{nn} are said to be diagonal entries of matrix A.

Triangular matrix : A square matrix $A = [a_{ij}]$ is called an *upper triangular matrix* if $a_{ij} = 0$ for all $i > j$. A is called *lower triangular matrix* if $a_{ij} = 0$ for all $i < j$. In other words, a square matrix is called *upper triangular* if all the entries below the main diagonal are zeros. Similarly, a square matrix is called *lower triangular* if all the entries above the main diagonal are zeros. A matrix is called *triangular matrix* if it is either upper or lower triangular.

For example, $A = \begin{bmatrix} -1 & 2 & -1 & 4 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

is 4×4 upper triangular matrix and $B = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 2 & 0 \\ -1 & 4 & 3 \end{bmatrix}$

is 3×3 lower triangular matrix.

Diagonal matrix : A square matrix in which all entries off a matrix rather than the main diagonal are zeros is called a *diagonal matrix.* If A is $n \times n$ diagonal matrix, then we write A as $A = \text{diag.}\ (a_{11}, a_{22}, ..., a_{nn})$.

Scalar matrix : A diagonal matrix in which all the entries on main diagonal are equal is called a *scalar matrix.*

Identity matrix : A scalar matrix in which all the entries on main diagonal are equal to 1 is called the *identity matrix.*

Example 11 : $\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix}$

are some examples of diagonal matrices.

$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}$

are some examples of scalar matrices.

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

are identity matrices of order 2, 3 and 4 respectively. The identity matrix of order n is usually denoted by I_n.

Transpose of a matrix : If $A = [a_{ij}]_{m \times n}$ is a matrix, then the *transpose of A* denoted by A^t, is defined to be a matrix obtained by converting rows of A into columns and columns into rows. Thus,

$$A^t = [a_{ji}]_{n \times m}$$

Note that if A is $m \times n$ matrix, then clearly A^t is $n \times m$ matrix, and the first row of A is the first column at A^t, second row of A is the second column A^t, and so on.

Example 12 :

(i) If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, then $A^t = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$

(ii) If $B = \begin{bmatrix} -1 & 2 \\ -2 & 4 \\ 4 & 0 \\ 1 & -1 \end{bmatrix}$ then $B^t = \begin{bmatrix} -1 & -2 & 4 & 1 \\ 2 & 4 & 0 & -1 \end{bmatrix}$

(iii) If $C = \begin{bmatrix} 2 & 0 & -5 & \frac{2}{3} & 6 \end{bmatrix}$, then $C^t = \begin{bmatrix} 2 \\ 0 \\ -5 \\ \frac{2}{3} \\ 6 \end{bmatrix}$

Symmetric Matrix : A square matrix $A = [a_{ij}]$ is said to be **symmetric** matrix if $A^t = A$; that is, $a_{ij} = a_{ji}$, for all i and j. A is said to be skew symmetric or antisymmetric if $A^t = -A$; that is $a_{ij} = -a_{ji}$ for all i and j.

For example, $\begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$, $\begin{bmatrix} a & b & c \\ b & g & h \\ c & h & f \end{bmatrix}$ and $\begin{bmatrix} 5 & -1 & 2 \\ -1 & 0 & 3 \\ 2 & 3 & -7 \end{bmatrix}$ are symmetric matrices.

$\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$ are skew symmetric matrices.

Note :

(1) In skew symmetric matrix, the entries on the main diagonal are zero.

(2) Given any square matrix A, the matrix $(A + A^t)$ is symmetric matrix and the matrix $A - A^t$ is skew symmetric. Using this we see that any square matrix can be written as a sum of symmetric matrix and a skew symmetric matrix, viz

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$$

(3) If A and B are symmetric matrices of the same size and if k is scalar, then A^t, A + B, A − B and kA are all symmetric. Here, it is worth to note that the product AB need not be symmetric.

(4) If A is invertible symmetric matrix, then A^{-1} is symmetric.

(5) The product AA^t and A^tA are symmetric.

Conjugate matrix : If $A = [a_{ij}]$ is a matrix with complex numbers as entries, the matrix $\bar{A} = [\bar{a}_{ij}]$ is called the complex conjugate matrix of A, where $\bar{a}_{ij}$ denotes complex conjugate of a_{ij}.

Adjoint matrix : If $A = [a_{ij}]$, $a_{ij} \in \mathbb{C}$, then the matrix $A^* = \bar{A}^t = [\bar{a}_{ij}]^t$ is called adjoint of A or complex conjugate transpose of A. Note : $(\bar{A})^t = (\bar{A}^t)$

For example, let $A = \begin{bmatrix} 1+i & 1 \\ 2+i & -i \end{bmatrix}$, then $\bar{A} = \begin{bmatrix} 1-i & 1 \\ 2-i & i \end{bmatrix}$ and $A^* = \begin{bmatrix} 1-i & 2-i \\ 1 & i \end{bmatrix}$.

We recall few definitions and results. If $Z = x + iy$ is complex number, then $\bar{Z} = x - iy$ is complex conjugate of Z. For complex numbers Z_1, Z_2, we know $\overline{Z_1 \pm Z_2} = \bar{Z}_1 \pm \bar{Z}_2$, $\overline{Z_1 Z_2} = \bar{Z}_1 \bar{Z}_2$, $(\bar{\bar{Z}}_1) = Z_1$, $|Z_1|^2 = Z_1\bar{Z}_1 \geq 0$ and $|Z_1| = 0$ if and only if $Z_1 = 0$. Also $|Z_1 + Z_2| \leq |Z_1| + |Z_2|$ and $|Z_1 Z_2| = |Z_1|\,|Z_2|$.

Using these facts to our recent definitions, we have

$$\overline{A+B} = \bar{A} + \bar{B}, \quad \overline{AB} = \bar{A}\,\bar{B}, \quad \bar{\bar{A}} = A.$$

And $(A+B)^* = A^* + B^*$, $(AB)^* = B^*A^*$ and $(A^*)^* = A$

If A is non-singular, then $(A^*)^{-1} = (A^{-1})^*$.

Unitary matrix : A unitary matrix is a complex square matrix A such that

$$AA^* = A^*A = I.$$

Orthogonal matrix : An orthogonal matrix is a real square matrix A such that

$$AA^t = A^tA = I$$

Note : If a matrix A is unitary, then $A^{-1} = A^*$.

and if A is orthogonal then $A^{-1} = A^t$.

The matrices $\begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & \frac{1}{2} \end{bmatrix}$, $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ are orthogonal matrices.

The matrices $\begin{bmatrix} -1/\sqrt{2} & -i/\sqrt{2} \\ i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$, $\begin{bmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1-i}{2} \end{bmatrix}$ are examples of unitary matrices.

Idempotent matrix : A square matrix A is said to be **idempotent** if AA = A; that is, $A^2 = A$.

For example, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 2 & 2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$ are idempotent matrices.

Note :

(i) The 2×2 matrix is idempotent if and only if it is diagonal with diagonal entries 1 or 0.

(ii) A matrix A is idempotent if and only if for all positive integers n, $A^n = A$.

(iii) Idempotent matrices arise frequently in regression analysis and econometrics.

1.3 Operations on Matrices

Definition : Two matrices are said to be *equal* only if they have the same order and the corresponding entries in the two matrices are equal.

For example, if

$$A = \begin{bmatrix} -1 & 3 \\ 7 & -8 \end{bmatrix} \quad B = \begin{bmatrix} 3 & -1 \\ 7 & -8 \end{bmatrix} \text{ and } C = \begin{bmatrix} 3 & -1 & 2 \\ 4 & 5 & 6 \end{bmatrix}$$

then $A \neq B$ since the *corresponding* entries are not equal. $A \neq C$, since A and C are not of the same order. Similarly $B \neq C$.

Definition : If A and B are any two matrices of the same order, then the *sum A + B* is a matrix obtained by adding together the corresponding entries in the two matrices.

For example, if

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 3 & 5 \\ -3 & 4 & 6 \\ 1 & 2 & 4 \end{bmatrix} \quad B = \begin{bmatrix} -2 & 4 & 6 \\ 6 & 0 & -1 \\ 2 & 3 & -3 \\ 1 & -1 & 1 \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 2 & 6 \end{bmatrix}$$

then A + B is defined and

$$A + B = \begin{bmatrix} -1 & 3 & 6 \\ 8 & 3 & 4 \\ -1 & 7 & 3 \\ 2 & 1 & 5 \end{bmatrix}$$

However, A + C or B + C cannot be defined as they are not of the same order.

Note : If $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{m \times n}$, then $A + B = [a_{ij} + b_{ij}]_{m \times n}$.

Definition : If A is any matrix and k is any scalar (real number), then the product kA is defined to be a matrix obtained by multiplying each entry of A by k.

Note : The product kA is called *scalar multiplication.*

Example 13 : If $A = \begin{bmatrix} 5 & -5 & 0 \\ 10 & 15 & 25 \\ -15 & 20 & 30 \end{bmatrix}$ then $\left(\frac{1}{5}\right) A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 3 & 5 \\ -3 & 4 & 6 \end{bmatrix}$

$\left(\frac{3}{5}\right) A = \begin{bmatrix} 3 & -3 & 0 \\ 6 & 9 & 15 \\ -9 & 12 & 18 \end{bmatrix}$, $(-1) A = \begin{bmatrix} -5 & 5 & 0 \\ -10 & -15 & -25 \\ 15 & -20 & -30 \end{bmatrix}$

Definition : If A and B are any two matrices, then the *subtraction (difference)* A – B is defined by

$$A - B = A + (-1) B$$

For example, if

$$A = \begin{bmatrix} -1 & 4 & 3 \\ 1 & -3 & 5 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -7 \end{bmatrix}$$

then $A - B = \begin{bmatrix} -2 & 6 & 2 \\ 1 & -5 & 12 \end{bmatrix}$

Note : If A is any matrix we define – A to be (– 1) A; that is – A = (– 1) A.

Definition : If A is m × n matrix and B is n × p matrix, then the product AB of A and B is the m × p matrix whose entries are determined as follows. To find the entry in i^{th} row and j^{th} column of AB, single out i^{th} row of A and j^{th} column of B. Multiply the corresponding entries from row and column together and add up the resulting products.

Example 14 : Consider the matrices

$$A = \begin{bmatrix} 4 & 1 & 2 \\ 0 & 6 & -2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 4 & 3 & 4 \\ -1 & 0 & -3 \\ 2 & 5 & 8 \end{bmatrix}$$

Since A is 2 × 3 matrix and B is 3 × 3 matrix, the product AB is a 2 × 3 matrix.

To determine, for example, the entry in second row and third column of AB, we single out second row from A and third column from B. Then, as illustrated below, we multiply corresponding entries together and add-up these products.

$$0 \times 4 + 6 \times (-3) + (-2) \times 8 = -34.$$

Thus, pictorally,

$$\begin{bmatrix} 4 & 1 & 2 \\ \boxed{0} & \boxed{6} & \boxed{-2} \end{bmatrix} \begin{bmatrix} 4 & 3 & \boxed{4} \\ -1 & 0 & \boxed{-3} \\ 2 & 5 & \boxed{8} \end{bmatrix} = \begin{bmatrix} \square & \square & \square \\ \square & \square & \boxed{-34} \end{bmatrix}$$

Computing in this way, we get

$$AB = \begin{bmatrix} 19 & 22 & 29 \\ -10 & -10 & -34 \end{bmatrix}$$

If $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{n \times p}$ are general matrices, then the entry in i^{th} row and j^{th} column of AB is given by

$$a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + ... + a_{in} b_{nj} = \sum_{k=1}^{n} a_{ik}\, b_{kj}$$

This fact is visualised in the following :

$$AB = \begin{bmatrix} a_{11} & a_{12} & ..... & a_{1n} \\ a_{21} & a_{22} & ..... & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & ..... & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & ..... & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & ..... & b_{1j} & ..... & b_{1p} \\ b_{21} & b_{22} & ..... & b_{2j} & ..... & b_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{n1} & b_{n2} & ..... & b_{nj} & ..... & b_{np} \end{bmatrix}$$

Morever, we can write

$AB = \left[\sum_{k=1}^{n} a_{ik}\, b_{kj} \right]$, where $i = 1, 2, ..., m$ and $j = 1, 2, ..., p$. Thus AB is of order $m \times p$.

Definition : A matrix, all of whose entries are zero, is called a *zero matrix* .

For example,

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

are zero matrices of order 2×2, 3×4, 3×3 and 4×1 respectively.

In practice, zero matrix is simply denoted by O. The order of zero matrix is assumed in such a way that the operations for which it is used are valid.

For example, if

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix}, \text{ then } A + O = A$$

where O is assumed to be of order 2×2.

Properties of Matrix Operations :

The following two theorems reveal the properties of the matrix operations that we have seen above.

Theorem 1 : If $A = [a_{ij}]$, $B = [b_{ij}]$ and $C = [c_{ij}]$ are matrices of the same order $m \times n$, then the following rules hold good :

(a) $A + B = B + A$ (commutative law for sum)

(b) $(A + B) + C = A + (B + C)$ (Associative law for sum)

(c) $A + O = A$

(d) $A + (-A) = O$

(e) $k(A + B) = kA + kB$

(f) $(k + l) A = kA + lA$

(g) $(kl) A = k(lA)$

(h) $1A = A$

where, k and l are real (scalars) numbers.

Proof : The proofs are simple and straight forward, however we prove (a) and (f), the rest are left as an exercise.

Proof of (a) : We have

$$\begin{aligned} A + B &= [a_{ij} + b_{ij}]_{m \times n} && \text{by definition of sum} \\ &= [b_{ij} + a_{ij}]_{m \times n} && \text{since } a_{ij}, b_{ij} \text{ are reals} \\ &= [b_{ij}]_{m \times n} + [a_{ij}]_{m \times n} && \text{again by definition of sum} \\ &= B + A \end{aligned}$$

Proof of (f) : We have

$$\begin{aligned} (k + l) A &= (k + l)\,[a_{ij}]_{m \times n} \\ &= [(k + l)\, a_{ij}]_{m \times n} && \text{by definition of scalar multiplication} \\ &= [ka_{ij} + la_{ij}]_{m \times n} && \text{since } k, l, a_{ij} \text{ are reals} \\ &= [k\, a_{ij}]_{m \times n} + [la_{ij}]_{m \times n} && \text{by definition of sum} \\ &= k[a_{ij}]_{m \times n} + l[a_{ij}]_{m \times n} && \text{by definition of scalar multiplication} \\ &= kA + lA \end{aligned}$$

Theorem 2 : If A, B and C are matrices such that the indicated operations can be performed, then the following rules are valid.

(a) $A(BC) = (AB) C$ (Associative law for multiplication)

(b) $A(B + C) = AB + AC$

(c) $(A + B) C = AC + BC$ (Distributive laws)

(d) $\alpha(AB) = (\alpha A) B = A(\alpha B)$, where α is scalar.

(e) $AI = A$, where I is unit matrix of suitable order.

(f) $AO = O$.

Proof : We prove (a) and (d), the proofs of (b), (c), (e) and (f) are left as an exercise.

Proof of (a) : For the operations to hold, we must have $A = [a_{ij}]_{m \times n}$, $B = [b_{ij}]_{n \times p}$ and $C = [c_{ij}]_{p \times q}$. Then, we see that A (BC) and (AB) C are both of order $m \times q$. Also note that $(k, j)^{th}$ entry in BC is

$$\sum_{t=1}^{p} b_{kt}\, c_{tj},\ k = 1, 2, \ldots\ldots, n,\ j = 1, 2, \ldots., q$$

Also $(i, t)^{th}$ entry in the matrix AB is

$$\sum_{k=1}^{n} a_{ik}\, b_{kt}\ i = 1, 2, \ldots\ldots, m\ ;\ t = 1, 2, \ldots., p$$

Now the $(i, j)^{th}$ entry in A (BC) is

$$\sum_{k=1}^{n} a_{ik} \left[(k, j)^{th} \text{ entry in BC } \right]$$

$$= \sum_{k=1}^{n} a_{ik} \left(\sum_{t=1}^{p} b_{kt}\, c_{tj} \right)$$

Since multiplication of real numbers is distributive over addition and addition is commutative, we can change the order of summation on the right hand side. Hence, $(i, j)^{th}$ entry in A(BC)

$$= \sum_{t=1}^{p} \left[\sum_{k=1}^{n} (a_{ik}\, b_{kt}) \right] c_{tj}$$

$$= \sum_{t=1}^{p} [(i, t)^{th} \text{ entry in AB}]\, c_{tj}$$

$$= (i, j)^{th} \text{ entry in (AB) C.}$$

Thus, $(i, j)^{th}$ entry in A (BC) is equal to $(i, j)^{th}$ entry in (AB) C, for all $i = 1, 2, \ldots\ldots, m$ and $j = 1, 2, \ldots., q$. Hence A(BC) = (AB) C.

Proof of (d) : We have

$$\alpha(AB) = \alpha \left[\sum_{k=1}^{n} a_{ik}\, b_{kj} \right]_{m \times p}$$

$$= \left[\alpha \sum_{k=1}^{n} a_{ik}\, b_{kj} \right]_{m \times p} \qquad \text{by definition of scalar multiplication}$$

$$= \left[\sum_{k=1}^{n} \alpha\, (a_{ik}\, b_{kj}) \right]_{m \times p} \qquad \text{since multiplication distributes over addition in } \mathbb{R}$$

Now, using multiplication is associative and commutative, we have

$$\alpha (a_{ik}\, b_{kj}) = (\alpha\, a_{ik})\, b_{kj} = a_{ik}\, (\alpha\, b_{kj})$$

Hence, $\alpha (AB) = \left[\sum_{k=1}^{n} (\alpha\, a_{ik})\, b_{kj}\right]_{m \times p} = (\alpha A)\, B$

and $\alpha (AB) = \left[\sum_{k=1}^{n} a_{ik}\, (\alpha\, b_{kj})\right]_{m \times p} = A\, (\alpha B)$

Invertible (non-singular) matrices :

Definition : If A is square matrix, and if there exists a matrix B such that

$$AB = BA = I$$

then A is said to be *invertible* matrix and B is called *inverse* of A.

Example 15 : If $A = \begin{bmatrix} 3 & 8 \\ 2 & 5 \end{bmatrix}$, then we have $B = \begin{bmatrix} -5 & 8 \\ 2 & -3 \end{bmatrix}$ such that

$$AB = \begin{bmatrix} 3 & 8 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} -5 & 8 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and $$BA = \begin{bmatrix} -5 & 8 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} 3 & 8 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore, B is inverse of A and A is invertible matrix.

Remark : If A is invertible, then there is B such that AB = BA = I, so we see A and B are both square matrices of the same order.

Example 16 : Show that the matrix

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 4 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

is not invertible.

Solution : Suppose there is a matrix B such that

$$AB = BA = I_3$$

Observe that in the product AB, third row will consist entirely of zeros, whereas the third row of I_3 does not consist of zeros only.

Hence, $AB = I_3$ is not possible for any matrix B. Therefore, A is not invertible.

Theorem 3 : If a matrix A is invertible, then inverse of A is unique.

Proof : Suppose B and C are two matrices and both are inverses of A. Then, by definition of inverse

$$AB = BA = I \quad \ldots (i)$$

$$AC = CA = I \quad \ldots (ii)$$

Now, $B = BI = B\,(AC)$ $\because$ AC = I by (ii)

$= (BA)\,C$ associativity of matrix multiplication

$= IC$ $\because$ $BA = I$ from (i)

$= C$

Thus $B = C$, showing that A cannot have more than one inverse; that is, inverse of A, if it exists, is unique.

Notation : Since inverse of a matrix A, if it exists is unique, we denote inverse of A by A^{-1}.

Thus, $AA^{-1} = A^{-1}A = I$

Theorem 4 : If A is invertible matrix, then

(a) A^{-1} is invertible and $(A^{-1})^{-1} = A$

(b) For any non-zero scalar α, αA is invertible and

$$(\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$$

Proof : (a) Since A is invertible, we have

$$AA^{-1} = A^{-1}A = I,$$

Hence by the definition of invertibility, A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

(b) We have

$$(\alpha A)\left(\frac{1}{\alpha} A^{-1}\right) = \frac{1}{\alpha}(\alpha A)\,A^{-1} \quad \text{using (d) of theorem (2)}$$

$$= \frac{1}{\alpha}\,[\alpha\,(AA^{-1})] \quad \text{using (d) of theorem (2)}$$

$$= \frac{1}{\alpha}\,\alpha\,(AA^{-1}) \quad \text{using (g) of theorem (1)}$$

$$= 1\,(AA^{-1})$$

$$= AA^{-1} \quad \because\ 1A = A$$

$$= I$$

Similarly we can show that

$$\left(\frac{1}{\alpha} A^{-1}\right)(\alpha A) = I$$

Therefore, by definition $\frac{1}{\alpha} A^{-1}$ is inverse of α A, hence symbolically,

$$(\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$$

Theorem 5 : If A and B are invertible matrices of the same order, then AB is invertible and moreover

$$(AB)^{-1} = B^{-1} A^{-1} \text{ (Reversal law)}$$

Proof : Since A and B are invertible,

$$AA^{-1} = A^{-1}A = I$$

and $$BB^{-1} = B^{-1}B = I$$

Now consider, $$(AB)\left(B^{-1} A^{-1}\right) = A (B B^{-1}) A^{-1} \quad \text{by associativity}$$

$$= A (I) A^{-1} \quad \because BB^{-1} = I$$

$$= AA^{-1}$$

$$= I \quad \because AA^{-1} = I$$

Similarly, we can show that

$$(B^{-1} A^{-1}) (AB) = I$$

Therefore, $$(AB) (B^{-1} A^{-1}) = (B^{-1} A^{-1}) (AB) = I$$

Hence AB is invertible and moreover

$$(AB)^{-1} = B^{-1} A^{-1}$$

Remark : The above theorem 5 can be generalized. If $A_1, A_2, \ldots, A_n$ are invertible matrices and the product $A_1 A_2 \ldots A_n$ is compatible, then $(A_1 A_2 \ldots A_n)$ is invertible and

$$(A_1 A_2 \ldots A_n)^{-1} = A_n^{-1} A_{n-1}^{-1} \ldots A_2^{-1} A_1^{-1}$$

Theorem 6 : (a) For any matrix (A), $(A^t)^t = A$

(b) If A and B are matrices of the same order, then

$$(A + B)^t = A^t + B^t$$

(c) If α is any scalar and A is any matrix, then

$$(\alpha A)^t = \alpha A^t$$

(d) If A and B are any two matrices such that the product AB is defined, then

$$(AB)^t = B^t A^t$$

Proof : Proofs of (a), (b) and (c) are straight forward from the definition of transpose of a matrix, hence are left as an exercise. We prove here only (d).

Proof of (d) : Let $A = [a_{ij}]_{m \times n}$ and $B = [b_{ij}]_{n \times p}$. Then AB is defined and is of order $m \times p$. A^t and B^t are of order $n \times m$ and $p \times n$, hence product $B^t A^t$ is defined and is of order $p \times m$. Thus $(AB)^t$ and $B^t A^t$ are of the same order.

Now, from definition, $(i, j)^{th}$ entry of A = $(j, i)^{th}$ entry of A^t, similar arguments with B^t, $(AB)^t$ and $B^t A^t$, we have

$$(i, j)^{th} \text{ entry of } (AB)^t = (j, i)^{th} \text{ entry of } AB$$

$$= \sum_{k=1}^{n} a_{jk} b_{ki}$$

$$= \sum_{k=1}^{n} \left[(j, k)^{th} \text{ entry of A}\right] \left[(k, i)^{th} \text{ entry of B}\right]$$

$$= \sum_{k=1}^{n} [(k, j)^{th} \text{ entry of } A^t] [(i, k)^{th} \text{ entry of } B^t]$$

$$= \sum_{k=1}^{n} [(i, k)^{th} \text{ entry of } B^t] [(k, j)^{th} \text{ entry of } A^t]$$

since entries are reals

$$= (i, j)^{th} \text{ entry of } B^t A^t$$

This is true for each i = 1, 2, ..., p and j = 1, 2, ..., m. Hence $(AB)^t = B^t A^t$.

Remark : If $A_1, A_2, ..., A_n$ are matrices and the product $A_1, A_2 ..... A_n$ is compatible, then

$$(A_1 A_2 ..... A_n)^t = A_n^t \; A_{n-1}^t \; ..... A_2^t \; A_1^t$$

This is generalisation of (d) in theorem (6).

Trace of a matrix :

Let A be a square matrix of order n then trace of matrix A is denoted by tr(A) and it is defined as sum of the diagonal elements of matrix A.

Namely $\quad tr(A) = a_{11} + a_{22} + a_{33} + \ldots + a_{nn}$

Theorem : Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are n-square matrices and k is a scalar then

(i) $tr (A + B) = tr (A) + tr (B)$

(ii) $tr (A^T) = tr (A)$

(iii) $tr (kA) = k\, tr (A)$

(iv) $tr (BA) = tr (AB)$.

Note : All the proofs of the theorems are straight forward.

Example :

(1) Let A and B be the square matrices of order 3.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{bmatrix}$$

then $\quad tr (A) = \text{Sum of diagonal of matrix A}$

$$= 1 + (-4) + 7 \Rightarrow 4$$

$$tr (B) = \text{Sum of diagonals of matrix B}$$

$$= 2 + 3 - 4 \Rightarrow 1$$

Note : From the given matrices A and B the properties of trace of matrix (in the above theorem) are easily verified.

Adjoint of a square matrix :

Let A be any given square matrix.

Definition : The matrix $[A_{ij}]$, where A_{ij} denotes the cofactor of a_{ji} in the determinant |A|, is called the adjoint of A and is denoted by the symbols adj (A). So that the adjoint of A is the transpose of the matrix formed by the cofactors of A.

Thus the adjoint of any given square matrix. A is obtained by replacing every $(i, j)^{th}$ element of A by the cofactor of the $(j, i)^{th}$ element in |A|.

If $$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

then $$\text{adj } (A) = \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}$$

Thus to obtain the adjoint of a given matrix, first obtain the transpose of matrix A and replace each element of transpose of matrix A by it's cofactor.

Example : Let $A = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & -1 \\ 2 & 0 & -4 \end{bmatrix}$ find adj (A).

Solution : We have, $A' = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 3 & -1 & -4 \end{bmatrix}$ $\therefore$ Adj $(A) = \begin{bmatrix} 4 & 4 & -4 \\ -2 & -10 & 1 \\ -2 & 2 & 1 \end{bmatrix}$

Properties of adjoint of a matrix :

(i) $A \cdot \text{adj } (A) = |A| I = \text{adj } (A) \cdot A.$

(ii) If $|A| \neq 0$ then $A\left[\frac{1}{|A|} \text{adj } (A)\right] = \left[\frac{1}{|A|} \text{adj } (A)\right] A = I.$

In this case $\frac{1}{|A|}$ adj (A) is the inverse of A.

(iii) If I_n be a unit matrix of order n then adj $(I_n) = I_n$.

(iv) If A be a square matrix and A' is a transpose of a matrix A then adj (A') = [adj (A)]'.

(v) If A is a symmetric matrix then adj (A) is also symmetric matrix.

(vi) If A be $n \times n$ matrix then $|\text{adj } (A)| = |A|^{n-1}$, if $|A| \neq 0$.

(vii) If A and B are square matrices of the same order, then

$$\text{adj } (AB) = \text{adj } (B) \cdot \text{adj } (A)$$

Singular and Non-singular matrices :

Definition : A square matrix A is said to be non-singular if $|A| \neq 0$ otherwise it is said to be singular matrix.

Notation : If A is square matrix, |A| denotes the determinant of A.

1.4 Block Multiplication of Matrices

Definiton : If $A = [a_{ij}]_{m \times n}$ is a matrix, then a matrix obtained from A by deleting some rows and/or some columns is called a *submatrix* of A.

For example, if

$$A = \begin{bmatrix} 0 & -1 & 3 & -2 \\ 4 & 1 & -1 & 0 \\ -3 & 1 & 4 & 2 \end{bmatrix}$$

then $\begin{bmatrix} 0 & -1 & -2 \\ 4 & 1 & 0 \end{bmatrix}$

is submatrix of A obtained by deleting third row and third column and

$$\begin{bmatrix} 0 & -1 & 3 \\ 4 & 1 & -1 \\ -3 & 1 & 4 \end{bmatrix}$$

is also a submatrix of A obtained by deleting just fourth column.

Definition : Given a matrix A, we can obtain a set of submatrices of A by drawing horizontal and vertical lines in the rectangular array for A, then the set of submatrices of A obtained in this way is called a *partition* of A.

Example 18 : Let

$$A = \left[\begin{array}{c|cc|c} 0 & -1 & 3 & -2 \\ 4 & 1 & -1 & 0 \\ \hline -3 & 1 & 4 & 2 \end{array}\right]$$

be matrix and two vertical and one horizontal lines be drawn, so that we obtain

$$A = \left[\begin{array}{cc|c} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \end{array}\right]$$

where $P_{11} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}, P_{12} = \begin{bmatrix} -1 & 3 \\ 1 & -1 \end{bmatrix}, P_{13} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$

etc. is a partition of A.

Theorem 7 : If A and B are two matrices given by

$$A = \left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ \hline a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{array}\right] \quad B = \left[\begin{array}{c|ccc|cc} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} \\ \hline b_{31} & b_{32} & b_{33} & b_{34} & b_{35} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} \end{array}\right]$$

Let the partitions of A and B are obtained by drawing vertical and horizontal lines as shown above. Then,

$$A = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \\ P_{31} & P_{32} \end{bmatrix} \text{ and } B = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \end{bmatrix}$$

where, $P_{11} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, ..., $P_{32} = \begin{bmatrix} a_{53} & a_{54} \\ a_{63} & a_{64} \end{bmatrix}$

and $Q_{11} = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}$, ..., $Q_{23} = \begin{bmatrix} b_{34} & b_{35} \\ b_{44} & b_{45} \end{bmatrix}$

Then, $AB = \begin{bmatrix} P_{11}Q_{11} + P_{12}Q_{21} & P_{11}Q_{12} + P_{12}Q_{22} & P_{11}Q_{13} + P_{12}Q_{23} \\ P_{21}Q_{11} + P_{22}Q_{21} & P_{12}Q_{12} + P_{22}Q_{22} & P_{21}Q_{13} + P_{22}Q_{23} \\ P_{31}Q_{11} + P_{32}Q_{21} & P_{31}Q_{12} + P_{32}Q_{22} & P_{31}Q_{13} + P_{32}Q_{23} \end{bmatrix}$

This theorem is called theorem of *block multiplication of matrices.*

SOLVED EXAMPLES

Example 1.1 : Show by an example that multiplication of matrices is *not commutative.*

Solution : Let $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 \\ 4 & 1 \end{bmatrix}$

then $AB = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 8 & 2 \end{bmatrix}$

and $BA = \begin{bmatrix} 0 & 2 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 8 & 6 \end{bmatrix}$

Therefore $AB \neq BA$

Example 1.2 : If $A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}$

then find (a) AB (b) (2A) B (c) A (2B) and (d) Verify 2 (AB) = (2A) B = A (2B)

Solution : (a) $AB = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}$

(b) $(2A)\,B = \begin{bmatrix} 6 & 0 \\ -2 & 4 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 24 & -6 \\ -8 & 10 \\ 8 & 2 \end{bmatrix} = 2\begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}$

(c) $A\,(2B) = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 8 & -2 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 24 & -6 \\ -8 & 10 \\ 8 & 2 \end{bmatrix} = 2\begin{bmatrix} 12 & -3 \\ -4 & 5 \\ 4 & 1 \end{bmatrix}$

(d) From (a), (b) and (c) we see that

2 (AB) = (2A) B = A (2B) is verified.

Example 1.3 : If A is invertible matrix, then A^t is also invertible and $(A^t)^{-1} = (A^{-1})^t$

Solution : Since A is invertible, we have

$$AA^{-1} = A^{-1}A = I$$

$$\therefore \quad (AA^{-1})^t = (A^{-1}A)^t = I^t$$

This implies

$$A^t.(A^{-1})^t = (A^{-1})^t . A^t = I \qquad (\because I^t = I)$$

Therefore, by definition of inverse

$(A^{-1})^t$ is inverse of A^t, hence A^t is invertible and

$$(A^t)^{-1} = (A^{-1})^t$$

Example 1.4 : Show that any square matrix can be written as sum of a symmetric matrix and skew-symmetric matrix.

Solution : Let A be any square matrix.

Then, $$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t) \qquad \ldots (*)$$

$$\text{Now, } \left[\frac{1}{2}(A + A^t)\right]^t = \frac{1}{2}(A + A^t)^t$$

$$= \frac{1}{2}\left(A^t - (A^t)^t\right)$$

$$= \frac{1}{2}(A^t + A), \qquad \because (A^t)^t = A$$

$\therefore$ $\frac{1}{2}(A + A^t)$ is symmetric matrix.

$$\text{Also, } \left[\frac{1}{2}(A - A^t)\right]^t = \frac{1}{2}(A - A^t)^t$$

$$= \frac{1}{2}\left(A^t - (A^t)^t\right)$$

$$= \frac{1}{2}(A^t - A)$$

$$= -\frac{1}{2}(A - A^t)$$

showing that $\frac{1}{2}(A - A^t)$ is antisymmetric matrix. Thus, (*) shows that A is a sum of symmetric matrix and skew-symmetric matrix.

Example 1.5 : If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is any 2×2 matrix, then show that if A is invertible, then $ad - bc \neq 0$ and in this case

$$A^{-1} = \begin{bmatrix} \frac{d}{ad - bc} & \frac{-b}{ad - bc} \\ \frac{-c}{ad - bc} & \frac{a}{ad - bc} \end{bmatrix}$$

Solution : Suppose A has the inverse

$$B = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$$

Then $$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hence,

$$\left.\begin{aligned} ax + bz &= 1 \\ ay + bw &= 0 \\ cx + dz &= 0 \\ cy + dw &= 1 \end{aligned}\right\} \quad \ldots (1)$$

Multiplying the first and third equations by d and – b respectively, and adding, we get

$$(ad - bc)\, x = d$$

Since x, y, z and w exist, this shows that we must have

$$ad - bc \neq 0$$

Then, $x = \dfrac{d}{ad - bc}$, and using this in third equation we get

$$z = \frac{-c}{ad - bc}$$

Similarly, from the second and fourth equations in (1), we obtain

$$y = \frac{-b}{ad - bc} \text{ and } w = \frac{a}{ad - bc}$$

Thus, $$A^{-1} = \begin{bmatrix} \dfrac{d}{ad - bc} & \dfrac{-b}{ad - bc} \\ \dfrac{-c}{ad - bc} & \dfrac{a}{ad - bc} \end{bmatrix}$$

Example 1.6 : Solve the following system of linear equations by variable elimination method.

$$x + y + z = 1 \quad \ldots \text{(i)}$$

$$2x - y + z = 2 \quad \ldots \text{(ii)}$$

Solution : Adding (i) and (ii) we get new equivalent system

$$x + y + z = 1$$

$$3x + 2z = 3$$

From these equations we can write

$$x = 1 - \frac{2}{3} z$$

and $$y = 1 - z - x = 1 - z - \left(1 - \frac{2}{3} z\right) = \frac{-z}{3}$$

Thus, assigning arbitrary value to z, we obtain

$$x = 1 - \frac{2}{3}\, t$$

$$y = -\frac{1}{3}\, t$$

$$z = t \,, t \in \mathbb{R}$$

is a set of solution of the given system.

This set is sometimes written as :

$$S = \left\{ \left(1 - \frac{2}{3}t, \frac{-1}{3}t, t\right) \mid t \in \mathbb{R} \right\}$$

or $$S = (1, 0, 0) + \left\{ t\left(-\frac{2}{3}, \frac{-1}{3}, 1\right) \mid t \in \mathbb{R} \right\}$$

Example 1.7 : Let $A = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline a & b & -1 & 0 \\ c & d & 0 & -1 \end{array}\right]$

Use the given partition to prove that $A^2 = I$, hence find A^{-1}.

Solution : We have

$$A = \left[\begin{array}{c|c} I_2 & O \\ \hline P & -I_2 \end{array}\right]$$

where $$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Using block multiplication, we have,

$$A^2 = \left[\begin{array}{c|c} I_2 & O \\ \hline P & -I_2 \end{array}\right] \left[\begin{array}{c|c} I_2 & O \\ \hline P & -I_2 \end{array}\right] = \left[\begin{array}{c|c} I_2 & O \\ \hline P - P & -I_2 \end{array}\right]$$

$$= I_4$$

As $A^2 = I_4$, A.A. $= I_4$, hence A is invertible and $A^{-1} = A$; by definition of inverse.

Example 1.8 : Let A and B be $n \times n$ matrices, either both symmetric or both antisymmetric. Prove that AB is symmetric if and only if AB = BA.

Solution : Case (i) : Suppose both A and B are symmetric, then $A^t = A$ and $B^t = B$.

If AB is symmetric, then $(AB)^t = AB$.

But $(AB)^t = B^t A^t = BA$, since $A^t = A$ and $B^t = B$.

$\therefore$ $AB = BA$

Conversely, suppose AB = BA

Now, $(AB)^t = B^t A^t$

$= BA$

$=$ AB by assumption.

$\therefore$ AB is symmetric.

Case (ii) : Suppose both A and B are skew-symmetric, so that $A^t = -A$ and $B^t = -B$.

If AB is symmetric, we have

$$(AB)^t = AB$$

But $(AB)^t = B^t A^t = (-B)(-A) = BA$

Thus, $(AB)^t = BA$

Hence $AB = BA$

Conservely, suppose $AB = BA$

Now, $(AB)^t = B^t A^t$

$= (-B)(-A)$ by assumption

$= BA$

$= AB$ by assumption

$\therefore$ AB is symmetric.

Exercise 1.1

1. Solve the following systems of linear equations and write down the set of solutions in each case.

(a) $\begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 8z &= 4 \\ 3x + 2y + 17z &= 1 \end{aligned}$

(b) $\begin{aligned} x_1 - 2x_2 + x_3 + x_4 &= 1 \\ x_1 - 2x_2 + x_3 + 5x_4 &= 5 \\ x_1 - 2x_2 + x_3 - x_4 &= -1 \end{aligned}$

(c) $\begin{aligned} x + 4z - w &= 7 \\ y - 2z - 3w &= 8 \end{aligned}$

(d) $\begin{aligned} x - y &= 0 \\ 3x + 4y &= 0 \end{aligned}$

(e) $\begin{aligned} 3x_1 + 4x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$

2. State true or false in each of the following and justify your answer :

(a) $\sqrt{2}\ x - \pi y + 4\sqrt{2}\ z = \sqrt{5}$ is a linear equation.

(b) $x^2 + 3y + z = 6$ is a linear equation.

(c) $x + y = 2$ has solution $x = -98$, $y = 100$

(d) $x + y = 0$ is a homogeneous system of linear equations

(e) $x - y - 3 = 0$
$x + 3y + 2z + 10 = 0$
is homogeneous system of linear equations.

(f) The equation $x_1 + 5x_2 - 3x_3 = -1$ has no solutions.

(g) $2x + 3y = 1$
$10x + 15y = 30$
has a solution.

(h) $x - y = 5$
$3x + 4y = 1$
has unique solution, that is only one solution.

(i) Homogeneous system of linear equations is always consistent.

(j) A non-homogeneous system of linear equations is always consistent.

3. State true or false for each of the following and justify your answer :
   (a) If matrix A is of order 2×3 and matrix B is of order 3×4 then A + B is defined.
   (b) If $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 5 & 0 \end{bmatrix}$, then A = B
   (c) A – B = A + (– 1) B
   (d) Triangular matrix is necessarily a square matrix.
   (e) For any two matrices A and B of the same order, AB is defined.
   (f) If for matrices A and B, AB is defined, then BA is also defined.
   (g) If for two matrices A and B if AB and BA are both defined, then both AB and BA are square matrices.
   (h) If for two matrices A and B, both AB and BA are defined, then AB = BA.
   (i) The product of any two triangular matrices is a triangular matrix.
   (j) If A is a matrix of order 3×4, and
   $$A + O = A$$ is true then O must be a zero matrix of order 3×4
   (k) If A is a matrix of order 3×4 and I is unit matrix such that IA is defined, then I is unit matrix of order 3×3.
   (*l*) If $A = \begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix}$, then A is invertible.
   (m) If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ then
   $$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$
   (n) If A is symmetric matrix then $A + A^t$ and AA^t are both symmetric.
   (o) The sum of two invertible matrices is invertible.
   (p) If A and B are non-singular matrices and AB = 0, then atleast A = 0 or B = 0.
   (q) If A and B are square matrices such that AB = 0, then it is possible that $A \neq 0$ and $B \neq 0$.
4. Show that the product of two diagonal matrices is again a diagonal matrix.
5. Show that if A has a row of zeros and B is any matrix for which AB is defined, then AB also has a row of zeros.
6. Show that product of two triangular matrices is also a triangular matrix.

7. If A is a scalar matrix of order $n \times n$, then show that A commutes with every $n \times n$ matrix.
8. Show that if $A = \text{diag.}\ (d_1, d_2, ..., d_n)$ and $d_1, d_2, ..., d_n$ are all non-zero, then
$$A^{-1} = \text{diag.}\left(\frac{1}{d_1}, \frac{1}{d_2}, ... , \frac{1}{d_n}\right)$$
9. Show that a square matrix A can be uniquely expressed as A = B + C, where B is symmetric and C is skew-symmetric.
10. Let A and B be $n \times n$ matrices, either both symmetric or both skew-symmetric. Prove that AB is symmetric if and only if AB = BA.
11. Solve the following matrix equation :
$$\begin{bmatrix} x+y & x-z \\ 3z+w & 2x-4w \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & -8 \end{bmatrix}$$
12. Consider the matrices
$$A = \begin{bmatrix} 2 & 3 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 3 \\ -2 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 4 & 1 \\ -1 & 0 & 5 \end{bmatrix}$$
$$D = \begin{bmatrix} 0 & -1 & 1 \\ 5 & 2 & 1 \\ 2 & 3 & 4 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 3 \\ 6 & -1 & 4 \end{bmatrix}$$
Compute the following wherever possible
(a) – 3 (D + 2E), (b) 2 B – C, (c) AB, (d) BA, (c) (3E) D, (f) (AB) C.
13. Let A, B, C, D and E be matrices as given in exercise (12) above. Then verify the following rules :
(a) $A(B + C) = AB + AC$
(b) $(AB) C = A (BC)$
(c) $(\alpha + \beta) D = \alpha D + \beta D$
(d) $(\alpha\beta) E = \alpha(\beta E)$
(e) $\alpha(AB) = (\alpha A) B = A (\alpha B)$, where $\alpha = -2, \beta = 5$.
14. If A is invertible matix whose inverse is given by $\begin{bmatrix} 1 & 3 \\ -2 & -5 \end{bmatrix}$, then find A.
15. If A is invertible matrix and suppose that the inverse of 7A is
$$\begin{bmatrix} 3 & -7 \\ -1 & 2 \end{bmatrix}$$
then find the matrix A.

16. If $A = \begin{bmatrix} a_{11} & a_{12} & ..... & a_{1n} \\ a_{21} & a_{22} & ..... & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ a_{n1} & a_{n2} & ..... & a_{nn} \end{bmatrix}$

and $D = \begin{bmatrix} d_1 & 0 & ..... & 0 \\ 0 & d_2 & ..... & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & ..... & d_n \end{bmatrix}$

then find AD.

17. Consider D and E matrices as in exercise (12), then show that
(a) $(D^t)^t = D$, (b) $(D + E)^t = D^t + E^t$, (c) $(AB)^t = B^t A^t$, (d) $[(-3)D]^t = (-3)D^t$.

18. If $A = \begin{bmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{bmatrix}$ $B = \begin{bmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{bmatrix}$

then show that
(a) $AB = BA$
(b) $A^2 = \begin{bmatrix} \cos 2x & -\sin 2x \\ \sin 2x & \cos 2x \end{bmatrix}$

19. If $A = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}$ then show that $A^2 = A$.

20. Show that for any positive integer n,
if $A = \begin{bmatrix} x & 1 \\ o & x \end{bmatrix}$, $A^n = \begin{bmatrix} x^n & nx^{n-1} \\ 0 & x^n \end{bmatrix}$

21. Given $A = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline a & b & c & -1 & 0 & 0 \\ d & e & f & 0 & -1 & 0 \\ g & h & i & 0 & 0 & -1 \end{array}\right]$

Use the indicated partition to show that $A^2 = I_6$. Hence find A^{-1}.

Answers (1.1)

1. (a) $S = (-1, 2, 0) + \{t(-7, 2, 1) \mid t \in \mathbb{R}\}$

   (b) $S = (0, 0, 0, 1) + \{s(2, 1, 0, 0) + t(-1, 0, 1, 0) \mid s, t \in \mathbb{R}\}$

   (c) $S = (7, 8, 0, 0) + \{s(-4, 2, 1, 0) + t(1, 3, 0, 1) \mid s, t \in \mathbb{R}\}$

   (d) $S = \{(0, 0)\}$

   (e) $S = \{t(-3, 2, 1) \mid t \in \mathbb{R}\}$

2. T – for True and F – for False

   (a) T, (b) F, (c) T, (d) T, (e) F, (f) F, (g) F, (h) T, (i) T, (j) F.

3. (a) F, (b) F, (c) T, (d) T, (e) F, (f) F, (g) T, (h) F, (i) T, (j) T, (k) T, (*l*) F, (m) T, (n) T, (o) F, (p) F, (q) T.

11. $x = 2, y = -1, z = 0, w = 3$.

12. (a) $\begin{bmatrix} -6 & -3 & -9 \\ -33 & -18 & -21 \\ -42 & -3 & -36 \end{bmatrix}$ (b) not possible

    (c) $\begin{bmatrix} -4 & 9 \\ 2 & -1 \\ -1 & 4 \end{bmatrix}$ (d) not possible

    (e) $\begin{bmatrix} 21 & 12 & 18 \\ 48 & 30 & 51 \\ 9 & 12 & 63 \end{bmatrix}$ (f) $\begin{bmatrix} -17 & -16 & 41 \\ 5 & 8 & -3 \\ -6 & -4 & 19 \end{bmatrix}$

14. $A = \begin{bmatrix} -5 & -3 \\ 2 & 1 \end{bmatrix}$

15. $\begin{bmatrix} -\frac{2}{7} & 1 \\ \frac{1}{7} & -\frac{3}{7} \end{bmatrix}$

16. $\begin{bmatrix} d_1a_{11} & d_2a_{12} & \ldots\ldots & d_na_{1n} \\ d_1a_{21} & d_2a_{22} & \ldots\ldots & d_na_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ d_1a_{n1} & d_2a_{n2} & \ldots\ldots & d_na_{nn} \end{bmatrix}$

1.5 Elementary Operations

Given a matrix A, we can perform some operations on rows and/or columns of A, and obtain a comparatively simple matrix, which enables us to solve some problems easily. In particular these operations have effective role in solving system of linear equations.

Definition : The following three operations on the rows of any given matrix are called *elementary row operations :*

(a) Multiplying all the entries of a row by a non-zero constant.

(b) Interchange of any two rows.

(c) Multiplying all the entries of any row by a constant and the resulting products to be added to the corresponding entries of any other row.

Notations : If the rows of a matrix are denoted by $R_1, R_2, ..., R_m$ from top to bottom, then the above three elementary row operations can be conveniently expressed in the following manner :

(i) kR_i, to mean that $k \neq 0$ and multiply all the entries of i^{th} row by a constant k.

(ii) R_{ij}, to mean that interchange of i^{th} row and j^{th} row.

(iii) $R_i + kR_j$, to mean that multiply all the entries of j^{th} row by a constant k and add the resulting products to the corresponding entries of i^{th} row.

Definition : Analogus to elementary row operations, we can define column operations, which are called *elementary column* operations as follows : If $C_1, C_2, ..., C_n$ are columns of a matrix from left hand side to right hand side, then we have

(a) kC_i, to mean that $k \neq 0$ and multiply i^{th} column by a constant k.

(b) C_{ij}, to mean the interchange of i^{th} column and j^{th} column.

(c) $C_i + kC_j$, to mean that multiply all the entries in j^{th} column by a constant k and add the resulting proudcts to the corresponding entries of i^{th} column.

Definition : Given a matrix A. By applying a finite sequence of *elementary row operations*, we obtain a simpler matrix A' :

$$A \xrightarrow{\text{seuence of row operation s}} \rightarrow ..... \rightarrow A' \quad \text{(A' is not transpose of matrix A).}$$

This process of obtaining A' is called *row-reduction, or Gaussian elimination.*

Example 1 : Let

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix}$$

be given a matrix. Obtain row-reduction of A.

Solution : We have

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 1 & 2 & 8 & 4 & 12 \end{bmatrix} \xrightarrow{R_3 - R_1}$$

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 2 & 6 & 3 & 7 \end{bmatrix} \xrightarrow{R_3 + (-2) R_2} \begin{bmatrix} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

$$\xrightarrow{R_2 - R_1} \xrightarrow{R_1 - R_3} \begin{bmatrix} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}$$

Definition : A matrix is said to be in *reduced row-echelon form* if it has the following properties :

1. If a row does not consist entirely of zeros, then the first non-zero number in the row is a 1. (We call this 1 a *leading 1*).
2. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
4. Each column that contains a leading 1 has zeros everywhere else.

Definition : A matrix having only the above three properties 1, 2 and 3 is called to be *row-echelon form.*

Example 2 : The zero matrix, the identity matrix of any order, are trivial examples of reduced row-echelon form. The matrices

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 6 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

are some more examples of reduced-row echelon form.

$$\begin{bmatrix} 1 & 4 & 3 & 7 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are examples of row-echelon forms.

Note : Clearly, reduced row-echelon form is always row-echelon form but not vice-versa.

1.6 Applications of Row-reduction for Solving Systems of Linear Equations

Given a homogeneous or non-homogenous system of linear equations, we consider its augmented matrix and reduce it to row-echelon form or reduced row-echelon form. The system of linear equations corresponding to the row-echelon form or reduced row-echelon form is simpler to solve. The original system of linear equations and the system corresponding to row-echelon form or reduced row-echelon form have the same solutions.

Definition : Given a system of linear equations, if we reduce the augmented matrix of this system to row-echelon form, then write the corresponding system and solve it by back substitution. This method of solving the given system of linear equation is known as *Gaussian elimination* method. If the augmented matrix is reduced to reduced row-echelon form, then write the corresponding system of linear equation and solve it. This method of solving a system of linear equations is known as Gauss-Jordan elimination method.

Note : Gauss-Jordan elimination method is easier to solve the system than Gaussian elimination. Moreover, the first method leads closure to the solution than the later one.

We illustrate these methods in the following examples :

SOLVED EXAMPLE

Example 1.9 : Consider the system of linear equations

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 5 \\ x_1 + x_2 + 5x_3 + 2x_4 &= 7 \\ x_1 + 2x_2 + 8x_3 + 4x_4 &= 12 \end{aligned}$$

Solution : Consider the augmented matrix of this system and apply operations. We have

$$[A \mid B] = \left[\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 5 \\ 1 & 1 & 5 & 2 & 7 \\ 1 & 2 & 8 & 4 & 12 \end{array}\right]$$

$$\xrightarrow[R_3 + (-1)R_2]{R_2 + (-1)R_1} \left[\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 2 & 6 & 3 & 7 \end{array}\right]$$

$$\xrightarrow{R_3 + (-2)R_2} \left[\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] \quad \ldots (*)$$

$$\xrightarrow[R_1 + (-1)R_3]{R_2 + (-1)R_3} \left[\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] \quad \ldots (**)$$

For *Gauss-elimination* we write down the system of equations corresponding to the row-echelon form (*). Then we get

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 5 \\ x_2 + 3x_3 + 1\, x_4 &= 2 \\ x_4 &= 3 \end{aligned}$$

Using $x_4 = 3$ in first two equations and expressing x_1 and x_2 in terms of constants and x_3, we get

$$x_1 = 5 - 3 - 2x_3 = 2 - 2x_3$$

and $$x_2 = 2 - 3 - 3x_3 = -1 - 3x_3$$

Thus, the solution set S of the system is given by assigning arbitrary value t to x_3 as

$$\begin{aligned} x_1 &= 2 - 2t \\ x_2 &= -1 - 3t \\ x_3 &= t \\ x_4 &= 3,\ t \in \mathbb{R} \end{aligned}$$

or $$\begin{aligned} S &= \{2 - 2t, -1 - 3t, t, 3 \mid t \in \mathbb{R}\} \\ &= \{(2, -1, 0, 3) + t(-2, -3, 1, 0) \mid t \in \mathbb{R}\} \end{aligned}$$

For *Gauss-Jordan elimination method* we write down the system of equations corresponding to the reduced row-echelon form (**). Then we get

$$\begin{aligned} x_1 + 2x_3 &= 2 \\ x_2 + 3x_3 &= -1 \\ x_4 &= 3 \end{aligned}$$

This straight forward gives us

$$\begin{aligned} x_1 &= 2 - 2x_3 \\ x_2 &= -1 - 3x_3 \\ x_4 &= 3 \end{aligned}$$

Assigning arbitrary value t to x_3 we get the solution set S as

$x_1 = 2 - 2t$, $x_2 = -1 - 3t$, $x_3 = t$, $x_4 = 3$, $t \in \mathbb{R}$.

$$\begin{aligned} S &= \{2 - 2t, -1 - 3t, t, 3 \mid t \in \mathbb{R}\} \\ &= \{(2, -1, 0, 3) + t(-2, -3, 1, 0) \mid t \in \mathbb{R}\} \end{aligned}$$

We observe that the solution set for the system by both the methods – Gauss-elimination and Gauss-Jordan elimination – is the same. Also later method leads closure to the solution than the earlier method.

Theorem 7 : If A is a $n \times n$ invertible matrix, then system of linear equations $AX = B$ has a unique solution given by $X = A^{-1} B$.

Proof : Since $A(A^{-1}B) = (A\, A^{-1})\, B = IB = B$, $X = A^{-1} B$ is a solution.

Conversely suppose X_o is solution of the system $AX = B$, then

$$AX_o = B$$

Multiplying both sides of this equation by A^{-1} on the left we get

$$\begin{aligned} A^{-1}(AX_o) &= A^{-1}B \\ (A^{-1}A)X_o &= A^{-1}B \\ I X_o &= A^{-1}B \\ \therefore \quad X_o &= A^{-1}B \end{aligned}$$

Thus, $X = A^{-1}B$ is unique solution of the given system.

Corollary : If A is invertible $n \times n$ matrix, then the homogeneous system $AX = 0$ has no non-trivial solution.

Proof : Since A is invertible, $AX = 0$ has unique solution $A^{-1}O = O$. Hence $AX = O$ has only trivial (zero) solution, which is equivalent that the given system has no non-trivial solution.

Definition : If a matrix B is obtained from a matrix A by applying a sequence of elementary operations on a matrix A then A is called *row-equivalent* to B.

Theorem 8 : If A is an $n \times n$ matrix and the homogeneous system $AX = 0$ has only the trivial solution, then A must be row-equivalent to I_n.

Proof : Let $AX = O$ be the matrix form of the system

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n = 0 \\ \quad . \qquad\quad . \qquad\quad . \qquad . \\ \quad . \qquad\quad . \qquad\quad . \qquad . \\ \quad . \qquad\quad . \qquad\quad . \qquad . \\ a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + a_{nn}x_n = 0 \end{array}\right\} \quad \ldots (1)$$

and assume that the system has only the trivial solution. If we solve this system by Gauss-Jordan elimination, then the system of equations corresponding to the reduced row-echelon form of the augmented matrix will be

$$\left.\begin{array}{l} x_1 + 0x_2 + \ldots\ldots + 0x_n = 0 \\ 0x_1 + x_2 + \ldots\ldots + 0x_n = 0 \\ \quad . \qquad\quad . \qquad\quad . \qquad . \\ \quad . \qquad\quad . \qquad\quad . \qquad . \\ \quad . \qquad\quad . \qquad\quad . \qquad . \\ 0x_1 + 0x_2 + \ldots\ldots + x_n = 0 \end{array}\right\} \quad \ldots (2)$$

Thus, this means the augmented matrix

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots\ldots & a_{1n} & 0 \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{n1} & a_{n2} & \ldots\ldots & a_{nn} & 0 \end{array}\right]$$

for the system (1) can be reduced to the augmented matrix

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & \ldots\ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots\ldots & 0 & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & 0 & \ldots\ldots & 1 & 0 \end{array}\right]$$

for the system (2) by a sequence of elementary row operations. If we disregard the last column of zeros in each of these matrices, we can conclude that A can be reduced to I_n by a sequence of elementary row operations, that is A is row equivalent to I_n.

Theorem 9 : Let the system of linear equations AX = B have a solution X_o. Then the general solution of the system is $X = X_o + X'$, where X' is any solution of the homogeneous system AX = 0.

Proof : Since $X = X_o$ is a solution of AX = B, we have $AX_o = B$. Hence,

$$\begin{aligned} A(X_o + X') &= AX_o + AX' \\ &= B + O \quad \because AX' = O \\ \therefore A(X_o + X') &= B \end{aligned}$$

So, $X_o + X'$ is a solution of the system AX = B.

Now, suppose X = Y is any solution of AX = B. Then, AY = B and given that $AX_o = B$.

Hence, $A(Y - X_o) = O$

Therefore, $X = Y - X_o$ is a solution of the system AX = O. Now, writing $Y - X_o = X'$, we get $Y = X_o + X'$, where X' is a solution of AX = O.

Thus, any solution of AX = B is of the form $Y = X_o + X'$, where X' is solution of AX = O.

SOLVED EXAMPLES

Example 1.10 : Consider the system

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 5 \\ x_1 + x_2 + 5x_3 + 2x_4 &= 7 \\ x_1 + 2x_2 + 8x_3 + 4x_4 &= 12 \end{aligned}$$

We have solved this system AX = B where, $B = \begin{bmatrix} 5 \\ 7 \\ 12 \end{bmatrix}$ in example (4), the solution set is given by

$$S = \{(2, -1, 0, 3) + t(-2, -3, 1, 0) \mid t \in \mathbb{R}\}$$

Observe that $X_o = (2, -1, 0, 3)$ is a solution of the given system and $X' = t(-2, -3, 1, 0) = (-2t, -3t, t, 0)$ is arbitrary (general) solution of the system AX = O. (Check !). Therefore, $X = X_o + X'$ is the general solution or gives the set of all solutions of the system.

Example 1.11 : Solve the following systems by (a) Gaussian elimination method (b) Gauss-Jordan elimination method, (c) Find particular solution of the system AX = B.

1. $$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 1 \\ 3x_1 + 4x_4 &= 1 \\ x_1 - 4x_2 - 2x_3 - 2x_4 &= 0 \end{aligned}$$

2. $$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 0 \\ 3x_1 + 4x_4 &= 2 \\ x_1 - 4x_2 - 2x_3 - 4x_4 &= 2 \end{aligned}$$

3. $$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 0 \\ 3x_1 + 4x_4 &= 0 \\ x_1 - 4x_2 - 2x_3 - 2x_4 &= 0 \end{aligned}$$

4. $$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 2 \\ x_1 + 2x_2 + x_3 &= -1 \\ 2x_1 + x_2 - 6x_3 &= 4 \end{aligned}$$

5. $$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 + x_2 - 6x_3 &= 0 \end{aligned}$$

Solution : (1) We have AX = B, where

$$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 0 & 0 & 4 \\ 1 & -4 & -2 & -2 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Consider the augmented matrix $[A \mid B]$ and reduce it to row-echelon form and reduced row echelon form.

We have, $[A \mid B] = \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 3 & 0 & 0 & 4 & 1 \\ 1 & -4 & -2 & -2 & 0 \end{array}\right]$

$$\xrightarrow[R_3 + (-1)R_1]{R_2 + (-3)R_1} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -6 & -3 & 1 & -2 \\ 0 & -6 & -3 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{R_3 + (-1)R_2} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & -6 & -3 & 1 & -2 \\ 0 & 0 & 0 & -2 & 1 \end{array}\right]$$

$$\xrightarrow[\left(-\frac{1}{2}\right)R_3]{\left(-\frac{1}{6}\right)R_2} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \end{array}\right] \quad \ldots (*)$$

$$\xrightarrow{R_1 + (-2)R_2} \left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{4}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \end{array}\right]$$

$$\xrightarrow[R_1 + \left(-\frac{4}{3}\right)R_3]{R_2 + \frac{1}{6}R_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \end{array}\right] \quad \ldots (**)$$

(a) For Gaussian elimination method we consider the row-echelon form of the matrix

$[A \mid B]$, here it is given by (*) and write down the corresponding system of linear equations as

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 1 \\ x_2 + \frac{1}{2}\, x_3 - \frac{1}{6} x_4 &= \frac{1}{3} \\ x_4 &= -\frac{1}{2} \end{aligned}$$

Solving these equations we get

$$x_1 = \frac{1}{3} - \frac{4}{3}\, x_4,$$

and $x_2 = \frac{1}{3} - \frac{1}{2}\, x_3 + \frac{1}{6}\, x_4$

Since, $x_4 = -\frac{1}{2}$, we get the solution set as

$$x_1 = 1,\ x_2 = \frac{1}{4} - \frac{1}{2}\, t,\ x_3 = t,\ x_4 = -\frac{1}{2},\ t \in \mathbb{R}$$

Thus, the general solution of the system is X, where

$$\begin{aligned} X &= \left(1, \frac{1}{4} - \frac{1}{2} t, t, -\frac{1}{2}\right), t \in \mathbb{R} \\ &= \left(1, \frac{1}{4}, 0, -\frac{1}{2}\right) + t\left(0, -\frac{1}{2}, 1, 0\right), t \in \mathbb{R} \end{aligned}$$

(b) For Gauss-Jordan elimination method, we consider the reduced row-echelon form given by (**), and writting the corresponding system we get,

$$\begin{aligned} x &= 1 \\ x_2 + \frac{1}{2} x_3 &= \frac{1}{4} \\ x_4 &= -\frac{1}{2} \end{aligned}$$

Hence the solution set (general solution) is

$$\begin{aligned} X &= \left(1, \frac{1}{4} - \frac{1}{2} t, t, -\frac{1}{2}\right) \\ &= \left(1, \frac{1}{4}, 0, -\frac{1}{2}\right) + t\left(0, -\frac{1}{2}, 1, 0\right) t \in \mathbb{R} \end{aligned}$$

(c) $\left(1, \frac{1}{4}, 0, -\frac{1}{2}\right)$ is a particular solution of AX = B and $t\left(0, -\frac{1}{2}, 1, 0\right)$ is arbitrary (general solution) of AX = 0.

2. The augmented matrix for this system is

$$[A \mid B] \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 3 & 0 & 0 & 4 & 2 \\ 1 & -4 & -2 & -2 & 2 \end{array}\right]$$

$$\xrightarrow[R_3 + (-1)\,R_1]{R_2 + (-3)\,R_1} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 2 \\ 0 & -6 & -3 & -1 & 2 \end{array}\right]$$

$$\xrightarrow{R_3 + (-1)\,R_2} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & -6 & -3 & 1 & 2 \\ 0 & 0 & 0 & -2 & 0 \end{array}\right]$$

$$\xrightarrow[\left(-\frac{1}{2}\right)R_3]{\left(-\frac{1}{6}\right)R_2} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \quad \ldots (*)$$

$$\xrightarrow{R_1 + (-1)\,R_2} \left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{4}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow[R_2 + \frac{1}{6}R_3]{R_1 + \left(-\frac{4}{3}\right)R_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & -\frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \quad \ldots (**)$$

(a) For Gauss elimination using (*), we get the solution set (the general solution) as

$x_1 = \frac{2}{3}$, $x_2 = -\frac{1}{3} - \frac{1}{2}$ t, $x_3 = t$, $x_4 = 0$, $t \in \mathbb{R}$

or $X = \left(\frac{2}{3}, -\frac{1}{3}, 0, 0\right) + t\left(0, -\frac{1}{2}t, t, 0\right)$, $t \in \mathbb{R}$

(b) For Gauss-Jordan elimination method using (**), we get directly

$x_1 = \frac{1}{3}$, $x_2 = -\frac{1}{3} - \frac{1}{2}$ t, $x_3 = t$, $x_4 = 0$.

(c) The particular solution of AX = B is $x_1 = \frac{2}{3}$, $x_2 = -\frac{1}{3}$, $x_3 = 0$, $x_4 = 0$

The set $X = \left(0, -\frac{1}{2}t, t, 0\right)$, $t \in \mathbb{R}$ is general solution of AX = 0.

3. Since A is same as in (1) and (2) and in augmented matrix last column is zero, the row-echelon form and reduced row-echelon form are :

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \quad \ldots (*)$$

and

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \quad \ldots (**)$$

Using (*) and (**) above we have

(a) Gauss-elimination

$x_1 = 0, x_2 = -\frac{1}{2}\, t, x_3 = t, x_4 = 0, t \in \mathbb{R}$

(b) Gauss-Jordan elimination

$x_1 = 0, x_2 = -\frac{1}{2} t, x_3 = t, x_4 = 0, t \in \mathbb{R}$

(c) $(0, -1, 2, 0)$ is particular solution of $AX = 0$.

4. We have

$$[A \mid B] = \left[\begin{array}{ccc|c} 2 & 3 & -1 & 2 \\ 1 & 2 & 1 & -1 \\ 2 & 1 & -6 & 4 \end{array}\right]$$

$$\xrightarrow{R_{12}} \left[\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 2 & 3 & -1 & 2 \\ 2 & 1 & -6 & 4 \end{array}\right]$$

$$\xrightarrow[R_3 + (-2) R_1]{R_2 + (-1) R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & -1 & -3 & 4 \\ 0 & -3 & -8 & 6 \end{array}\right]$$

$$\xrightarrow{(-1) R_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & 1 & 3 & -4 \\ 0 & -3 & -8 & 6 \end{array}\right]$$

$$\xrightarrow{R_3 + 3R_2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & -1 \\ 0 & 1 & 3 & -4 \\ 0 & 0 & 1 & -6 \end{array}\right] \quad \ldots (*)$$

$$\xrightarrow{R_1 + (-2)R_2} \left[\begin{array}{ccc|c} 1 & 0 & -5 & 7 \\ 0 & 1 & 3 & -4 \\ 0 & 0 & 1 & -6 \end{array}\right]$$

$$\xrightarrow[R_2 + (-3)R_3]{R_1 + 5R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -23 \\ 0 & 1 & 0 & 14 \\ 0 & 0 & 1 & -6 \end{array}\right] \quad \ldots (**)$$

(a) For Gauss-elimination using (*), we can solve the system by back substitution, and get $x_1 = -23$, $x_2 = 14$, $x_3 = -6$

(b) For Gauss-Jordan elimination method, using (**), we get direct solution

$x_1 = -23$, $x_2 = 14$, $x_3 = -6$.

(c) The system has unique solution (only one solution).

5. Since the matrix A is same as in (4), we must have only trivial solution for the system $AX = 0$, namely $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ is the only solution. Therefore, in all cases (a), (b) and (c), the solution is the same : $x_1 = 0$, $x_2 = 0$, $x_3 = 0$.

Example 1.12 : Solve the following system :

$$\begin{aligned} x - y + 2z - w &= -1 \\ 2x + y - 2z - 2w &= -2 \\ -x + 2y - 4z + w &= 1 \\ 3x - 3w &= -3 \end{aligned}$$

Solution : The augmented matrix for the system is

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{array}\right]$$

We apply row operations to reduce this matrix to row-echelon form and further to reduced row-echelon form. The sequence

$$\xrightarrow{R_2 + (-2)R_1}, \xrightarrow{R_3 + R_1}, \xrightarrow{R_4 + (-3)R_1}$$ of operations gives

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 0 & 3 & -6 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -6 & 0 & 0 \end{array}\right]$$

$\xrightarrow{\frac{1}{3}R_2}$, $\xrightarrow{R_3 + (-1)R_2}$ and $\xrightarrow{R_4 + (-3)R_2}$ gives

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \quad \ldots (*)$$

Further $\xrightarrow{R_1 + R_2}$ gives

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \quad \ldots (**)$$

(*) is in row-echelon form and (**) is in reduced row echelon form.

Case (i) : For solving the system by Gauss-elimination method we write down the system of equations corresponding to (*), as

$$\begin{aligned} x - y + 2z - w &= -1 \\ y - 2z &= 0 \end{aligned}$$

From second equation we get $y = 2z$ and using this in first we obtain

$$\begin{aligned} x &= y - 2z + w - 1 = 2z - 2z + w - 1 = w - 1 \\ y &= 2z \end{aligned}$$

or $x = w - 1$ and $y = 2z$

Assigning arbitrary values to z and w, we obtain general solution as :

$x = s - 1, y = 2t, z = t, w = s; s, t \in \mathbb{R}$.

The solution set of the system is

$(s - 1, 2t, t, s) = (-1, 0, 0, 0) + t(0, 2, 1, 0) + s(1, 0, 0, 1)$

Observe that $(-1, 0, 0, 0)$ is a particular solution of the given system and $t(0, 2, 1, 0) + s(1, 0, 0, 1)$ is the general solution of the corresponding homogeneous system of the given system.

Case (ii) : For Gauss-Jordan elimination we use (**) and obtain the solution easily

$x = s - 1, y = 2t, z = t, w = s; s, t \in \mathbb{R}$ as in case (i).

Example 1.13 : Solve the following homogeneous system by Gauss-Jordan elimination method.

$$\begin{aligned} x_1 + 4x_2 + 5x_3 + 6x_4 + 9x_5 &= 0 \\ 3x_1 - 2x_2 + x_3 + 4x_4 - x_5 &= 0 \\ -x_1 - x_3 - 2x_4 - x_5 &= 0 \\ 2x_1 + 3x_2 + 5x_3 + 7x_4 + 8x_5 &= 0 \end{aligned}$$

Solution : Consider the augmented matrix

$$\left[\begin{array}{ccccc|c} 1 & 4 & 5 & 6 & 9 & 0 \\ 3 & -2 & 1 & 4 & -1 & 0 \\ -1 & 0 & -1 & -2 & -1 & 0 \\ 2 & 3 & 5 & 7 & 8 & 0 \end{array}\right]$$

$$\xrightarrow{R_2 + (-3)R_1}, \xrightarrow{R_3 + R_1}, \xrightarrow{R_4 + (-2)R_1}$$

$$\left[\begin{array}{ccccc|c} 1 & 4 & 5 & 6 & 9 & 0 \\ 0 & -14 & -14 & -14 & -28 & 0 \\ 0 & 4 & 4 & 4 & 8 & 0 \\ 0 & -5 & -5 & -5 & -10 & 0 \end{array}\right]$$

$$\xrightarrow{\left(-\frac{1}{14}\right)R_2}, \xrightarrow{R_3 + (-4)R_2}, \xrightarrow{R_4 + 5R_2}$$

$$\left[\begin{array}{ccccc|c} 1 & 4 & 5 & 6 & 9 & 0 \\ 0 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_1 + (-4)R_2}$$

$$\left[\begin{array}{ccccc|c} 1 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

This is in reduced row-echelon form. Writting the corresponding system we get

$$x_1 + x_3 + 2x_4 + x_5 = 0$$

$$x_2 + x_3 + x_4 + 2x_5 = 0$$

Assigning arbitrary values to x_3, x_4 and x_5 we obtain the general solution as

$$x_1 = -r - 2s - t$$

$$x_2 = -r - s - 2t$$

$$x_3 = r$$

$$x_4 = s$$

$$x_5 = t,\ r, s, t \in \mathbb{R}$$

In order tuple form

$(-r, -2s, -t, -r-s-2t, r, s, t)$

$= r(-1, -1, 1, 0, 0) + s(-2, -1, 0, 1, 0) + t(-1, -2, 0, 0, 1)$

Example 1.14 : For which values of k will the following system have (i) No solution ? (ii) Exactly one solution ? (iii) Infinitely many solutions ?

$$\begin{aligned} 4x + y + (k^2 - 14)z &= k + 2 \\ x + 2y - 3z &= 4 \\ 3x - y + 5z &= 2 \end{aligned}$$

Solution : Consider the augmented matrix

$$\left[\begin{array}{ccc|c} 4 & 1 & k^2-14 & k+2 \\ 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \end{array}\right]$$

$\xrightarrow{R_{12}}$, $\xrightarrow{R_{23}}$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 3 & -1 & 5 & 2 \\ 4 & 1 & k^2-14 & k+2 \end{array}\right]$$

$\xrightarrow{R_2 + (-3)R_1}$, $\xrightarrow{R_3 + (-4)R_1}$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & -7 & 14 & -10 \\ 0 & -7 & k^2-2 & k-14 \end{array}\right]$$

$\xrightarrow{\left(-\frac{1}{7}\right)R_2}$, $\xrightarrow{R_3 + 7R_2}$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & -2 & \frac{10}{7} \\ 0 & 0 & k^2-16 & k-4 \end{array}\right] \quad \ldots (*)$$

This matrix is in row-echelon form. There are three cases according to $k^2 - 16$ is zero or not.

Case (i) : If k = 4, the last row in (*) consists entirely of zeros, that is,

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & -2 & \frac{10}{7} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

We write down the corresponding system as

$$x + 2y - 3z = 4$$
$$y - 2z = \frac{10}{7}$$

Assigning arbitrary value to x_3, we obtain a solution set

$$x = \frac{8}{7} - t,\ y = \frac{10}{7} + 2t,\ x_3 = t,\ t \in \mathbb{R}$$

This shows that in case *k = 4*, the system has *infinitely many solutions.*

Case (ii) : If $k = -4$, the matrix (*) becomes :

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & -2 & \frac{10}{7} \\ 0 & 0 & 0 & 8 \end{array}\right]$$

This third row of this matrix will correspond to the equation of the form.

$0.x + 0.y + 0.z = -8$; that is , $0 = -8$ which is impossible. Hence in this case the system has *no solution*.

Case (iii) : If $k^2 - 16 \neq 0$; that is, if $k \neq \pm 4$, we can divide third row of the matrix in (*) by $k^2 - 16$ and obtain

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & -2 & \frac{10}{7} \\ 0 & 0 & 1 & \frac{k-4}{k^2-16} \end{array}\right]$$

or

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 4 \\ 0 & 1 & -2 & \frac{10}{7} \\ 0 & 0 & 1 & \frac{1}{k+4} \end{array}\right]$$

Then, the corresponding system of equations for this matrix is

$$x + 2y - 3z = 4$$
$$y - 2z = \frac{10}{7}$$
$$z = \frac{1}{k+1}$$

Solving by back-substitution we obtain :

$$x_1 = \frac{8}{7} - \frac{1}{k+4},\ y = \frac{10}{7} + \frac{2}{k+4},\ z = \frac{1}{k+4}$$

Thus, in this case the system *has only one solution*.

1.7 Elementary Matrices

Definition : An $n \times n$ matrix is called as *elementary matrix* if it can be obtained from $n \times n$ identity matrix I_n by performing a single elementary row or column operation.

Example 1 : The following are four elementary matrices and the operations that proudce them.

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$ by performing $R_1 + 2R_2$ on I_2

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$ by performing R_{23} on I_4

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ by performing $(-2)\, R_2$ on I_3

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$$ by performing $R_3 + \left(-\frac{1}{2}\right) R_1$ on I_3.

Note : The identity matrix is itself an elementary matrix, because it can be thought to be obtained from identity matrix by multiplying any row by 1.

Note : Since there are three elementary row operations, the elementary matrix coresponding to the operation R_{ij} is denoted by E_{ij}, the elementary matrix corresponding to the operation kR_i is denoted by $E_i\,(k)$, the elementary matrix corresponding to the operation $R_i + kR_j$ is denoted by $E_{i + (k)\, j}$. Thus, in example (1), the first matrix is $E_{1 + (2)\, 2}$, the second is E_{23}, the third is $E_2\,(-2)$ and the fourth one is $E_3 + \left(-\frac{1}{2}\right) 1$.

Remark 1 : If an elementary row operation is applied to an identity matrix I to produce an elementary matrix E, then there is an elementary row operation that when applied to E produces I back again. For the three row operations which produce elementary matrix E by applying on I, the list of the corresponding elementary row operations those applied on E will give back I, is given below :

Row operation on I that produces E	Row operation on E that produces I back
R_{ij} – interchange of i^{th} row and j^{th} row	R_{ij}, the same operation
(k) R_i – multiplying i^{th} row by a non-zero constant k	$\left(\frac{1}{k}\right) R_i$
$R_i + (k)\, R_j$ – add k-multiples of j^{th} row to corresponding elements of row i.	$R_i + (-k)\, R_j$.

The operations on the right side of this table are called *inverse operations.*

Remark 2 : There is no distinction between the elementary matrices obtained by using an elementary row operation and those obtained by using an elementary column operation. If we denote by E^{ij}, $E^{(k)\,i}$, $E^{i\,+\,(k)\,j}$, the elementary matrices obtained by applying the corresponding elementary column operations respectively, then one can easily verity that

$$E^{ij} = E_{ij}$$

$$E^{(k)\,i} = E_{(k)\,i}$$

$$E^{i\,+\,(k)\,j} = E_{i\,+\,(k)\,j}$$

Therefore, we shall use only the symbols E_{ij}, $E_{(k)\,i}$ and $E_{i\,+\,(k)\,j}$ to denote elementary matrices obtained by applying whether row operations or column operations.

Example 2 : In each of the following an elementary row operation is applied to the 3×3 identity matrix to obtain an elementary matrix E, then E is restored to the identity matrix by applying the inverse row operation.

$$\underset{I_3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \xrightarrow{(3)\,R_2} \underset{E}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \xrightarrow{\left(\frac{1}{3}\right) R_2} \underset{I_3}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_{13}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \xrightarrow{R_{13}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 + 3R_3} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 + (-3)\,R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Let us now prove some important results relating to elementary operations on product of two matrices.

Theorem 10 : If A and B are $m \times n$ and $n \times p$ matrices respectively, and if e is an elementary row operation and e' an elementary column operation, then

(a) e(AB) = e(A) B, in other words, e (row operation) applied on the product AB is the same as e applied on prefactor A of AB.

(b) e'(AB) = Ae(B), in other words e' (column operation) applied on the product AB is the same as e' applied on the post factor B of AB.

Proof : (a) We prove the theorem separately for each of the three kinds of elementary row operations. Let $A = [a_{ij}]$ and $B = [b_{ij}]$.

Case (i) : Let $e = R_{st}$, the interchange of s^{th} row and t^{th} row, and let i = 1, 2, ..., m, j = 1, 2, ..., p.

If $i \neq s$ and $i \neq t$, we have

$$(i, j)^{th} \text{ entry of } e(AB) = (i, j)^{th} \text{ entry of } AB$$

$$= \sum_{k=1}^{n} a_{ik}\, b_{kj}$$

$$= \sum_{k=1}^{n} [(i, k)^{th} \text{ entry of } e(A)]\ b_{kj}$$

$$= (i, j)^{th} \text{ entry of } e(A)\, B.$$

If i = s, then

$$(s, j)^{th} \text{ entry of } e(AB) = (t, j)^{th} \text{ entry of } AB$$

$$= \sum_{k=1}^{n} a_{tk}\, b_{kj}$$

$$= \sum_{k=1}^{n} [(s, k)^{th} \text{ entry of } e(A)]\, b_{kj}$$

$$= (s, j)^{th} \text{ entry } e(A)B.$$

Lastly, if i = t, then

$$(t, j)^{th} \text{ entry of } e(AB) = (s, j)^{th} \text{ entry of } AB$$

$$= \sum_{k=1}^{n} a_{sk}\, b_{kj}$$

$$= \sum_{k=1}^{n} [(t, k)^{th} \text{ entry of } e(A)]\, b_{kj}$$

$$= (t, j)^{th} \text{ entry of } e(A)\, B$$

Thus, for all i = 1, 2, ..., m and j = 1, 2, ..., p,

$$(i, j)^{th} \text{ entry of } e(AB) = (i, j)^{th} \text{ entry of } e(A)\, B$$

Hence, $e(AB) = e(A)\, B$

Case (ii) : If $e = kR_s$ (Multiplying s^{th} row by k)

If $i \neq s$, then

$$(i, j)^{th} \text{ entry of } e(AB) = (i, j)^{th} \text{ entry of } AB$$

$$= \sum_{k=1}^{n} a_{ik}\, b_{kj}$$

$$= \sum_{k=1}^{n} [(i, k)^{th} \text{ entry of } e(A)]\ b_{kj}$$

$$= (i, j)^{th} \text{ entry of } e(A)\, B.$$

If i = s, then

$$\begin{aligned}(s, j)^{th} \text{ entry of } e(AB) &= k\,[(s, j)^{th} \text{ entry of } AB] \\ &= k \sum_{k=1}^{n} a_{sk}\, b_{kj} \\ &= \sum_{k=1}^{n} (ka_{sk})\, b_{kj} \\ &= \sum_{k=1}^{n} [(s, k)^{th} \text{ entry of } e(A)]\, b_{kj} \\ &= (s, j)^{th} \text{ entry of } e(A)B.\end{aligned}$$

Hence, $e(AB) = e(A)\,B$

Case (iii) : If $e = R_{s + k(t)}$

If $i \neq s$, then

$$\begin{aligned}(i, j)^{th} \text{ entry of } e(AB) &= (i, j)^{th} \text{ entry of } AB \\ &= \sum_{k=1}^{n} a_{ik}\, b_{kj} \\ &= \sum_{k=1}^{n} [(i, k)^{th} \text{ entry of } e(A)]\; b_{kj} \\ &= (i, j)^{th} \text{ entry of } e(A)\,B\end{aligned}$$

If i = s, then

$$\begin{aligned}(s, j)^{th} \text{ entry of } e(AB) &= (s, j)^{th} \text{ entry of } AB + k\,[(t, j)^{th} \text{ entry of } AB] \\ &= \sum_{k=1}^{n} a_{sk}\, b_{kj} + k \sum_{k=1}^{n} a_{tk}\, b_{kj} \\ &= \sum_{k=1}^{n} a_{sk}\, b_{kj} + \sum_{k=1}^{n} (ka_{tk})\, b_{kj} \\ &= \sum_{k=1}^{n} [a_{sk} + ka_{tk}]\, b_{kj} \\ &= \sum_{k=1}^{n} [(s, k)^{th} \text{ entry of } e(A)]\, b_{kj} \\ &= (s, j)^{th} \text{ entry of } e(A)\,B\end{aligned}$$

Hence, $e(AB) = e(A)\,B$

This completes the proof of part (a) of the theorem.

(b) The proof is similar to the proof for (a) and left as an exercise.

Theorem 11 : If the elementary matrix E results from performing a certain row operation e on I_m and if A is an $m \times n$ matrix, then the product EA is the matrix that results when this same row operation is performed on A.

Proof : Since $A = I_m A$, by theorem (1) (a) we have

$e(A) = e(I_m A) = e(I_m) A$

But $e(I_m) = E$

Hence $e(A) = EA$

Theorem 12 : If the elementary matrix E results from performing a certain column operation e' on I_n and if A is an $m \times n$ matrix, then the product AE is the matrix that results when this same column operation is performed on A.

Proof : Since $A = AI_n$, by (b) of theorem (1) above

We have $e'(A) = e'(AI_n) = Ae'(I_n)$

But $e'(I_n) = E$

Hence $e'(A) = AE$

Theorem 13 : Every elementary matrix is invertible, and the inverse is also an elementary matrix.

Proof : If E is an elementary matrix, then E results from performing some row e (or column) operation on I. Let E_o be the matrix that results when the inverse of this operation e^{-1} is performed on I. Then by theorem (2) and using the fact that inverse row operations cancel the effect of each other, we see that $E_oE = e^{-1}(I) E = e^{-1}(IE) = e^{-1}(E) = e^{-1}[e(I)] = I$. Hence, $E_oE = I$ and similarly $EE_o = I$.

Thus, elementary matrix E_o is the inverse of E and E is invertible.

Theorem 14 : If A is an $n \times n$ matrix and A is row equivalent to I_n, then A is invertible.

Proof : As A is row equivalent to I_n, so that A can be reduced to I_n by a finite sequence of elementary row operations. By theorem (12), each of these operations can be accomplished by multiplying on the left by an appropriate elementary matrix. Thus, we can find elementary matrices $E_1, E_2, ..., E_t$ such that

$$E_t.E_{t-1} ..... E_2 E_1 A = I_n \quad \ldots (1)$$

By theorem (14), $E_1, E_2, ..., E_t$ are invertible. Multiplying both sides of equation (1) on the left successively by $E_t^{-1}, E_{t-1}^{-1}, ..., E_2^{-1}, E_1^{-1}$ we obtain

$$A = E_1^{-1} E_2^{-1} .....E_t^{-1} I_n = E_1^{-1} E_2^{-1} .....E_k^{-1} \quad \ldots (2)$$

Thus (2) expresses A as a product of invertible matrices, we conclude that A is invertible.

Note : From theorem (15), we see that A is invertible and equation (2) shows that A can be expressed as product of elementary matrices, since inverse of an elementary matrix is again an elementary matrix. Thus, we conclude that *invertible matrices are products of elementary matrices.*

Theorem 15 : A square matrix A is invertible if and only if its row echelon form is invertible.

Proof : Suppose A is an $n \times n$ matrix and that A is invertible. By theorem (15) and note, A is a product of elementary matrices. Let E_1, E_2, ... E_t be elementary matrices such that

$$A = E_1.E_2 ..... E_t \quad \ldots (1)$$

Multiplying both sides of equation (1) on the left successively by E_1^{-1}, E_2^{-1}, ... E_t^{-1}, we obtain

$$E_t^{-1} ..... E_2^{-1} E_1^{-1} A = I_n \quad \ldots (2)$$

Since E_1^{-1}, E_2^{-1}, ..., E_t^{-1} are elementary matrices, each of these corresponds to some elementary row operation performed on A. By theorem (12), this is equivalent to the fact that a sequence of t elementary row operations performed on A reduces A to I_n. Thus, from equation (2) we conclude that A is row-equivalent to I_n. That is, row echelon form of A is I_n. As I_n is invertible, we conclude that row echelon form of A is invertible.

Conversely, suppose row echelon form of A is A' and that A' is invertible. As A' is row echelon form of A, so that A can be reduced to A' by a finite sequence of elementary row operations. By theorem (12) each of these operations can be accomplished by multiplying on the left by an appropriate elementary matrix. Thus, we can find elementary matrices E_1, E_2, ..., E_r such that

$$E_r E_{r-1} ... E_2 E_1 A = A' \quad \ldots (3)$$

Since by theorem (14), E_1, E_2, ... , E_r are invertible, multiplying both sides of equation (3) on the left successively by E_r^{-1}, E_{r-1}^{-1} .... E_2^{-1}, E_1^{-1}, we obtain

$$A = E_1^{-1} E_2^{-1} ..... E_r^{-1} A' \quad \ldots (4)$$

As A' is given invertible and E_1^{-1}, E_2^{-1}, ... E_r^{-1} are invertible matrices, we see that A is product of invertible matrices. Hence A is invertible.

Computation of the inverse of a matrix by using Gaussian elimination method

Suppose $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

is the invertible matrix of order 3×3.

Let $X = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}$

be the inverse of A. Then

$$AX = I_3$$

or $$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \ldots (1)$$

The matrix equation (1) is equivalent to three matrix equations :

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \ldots (2)$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \ldots (3)$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \ldots (4)$$

Equations (2), (3) and (4) are systems of linear equations in three unknowns, we apply Gaussian elimination method to find x_1, y_1, z_1; x_2, y_2, z_2 and x_3, y_3, z_3. In this way we find X, the inverse of A by using *Gaussian elimination method*.

While solving the systems (2), (3) and (4) by Gaussian elimination method, we need to perform row operations on A to reduce it to row-echelon form. Since A is same in all these three equations, we can solve these systems simultaneously.

Let us illustrate this method by the following example.

SOLVED EXAMPLES

Example 1.15 : Compute the inverse of A by Gaussian elimination method, where

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 2 & 4 \end{bmatrix}$$

Solution : We consider augmented matrix in the following form which corresponds to all the three systems given by (2), (3) and (4).

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 5 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 3 & 2 & 4 & 0 & 0 & 1 \end{array}\right]$$

Now we perform row operations on A so that it reduces to row echelon form. We apply the same row operations simultaneously on I_3.

In the first step we apply,

$$\xrightarrow{R_2 + (-2)R_1} \quad \xrightarrow{R_3 + (-3)R_1} \text{ to get}$$

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 5 & 1 & 0 & 0 \\ 0 & -5 & -9 & -2 & 1 & 0 \\ 0 & -7 & -11 & -3 & 0 & 1 \end{array}\right]$$

In second step we perform $\xrightarrow{\left(-\frac{1}{5}\right)R_2}$ then $\xrightarrow{R_3 + 7R_2}$ to get

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 5 & 1 & 0 & 0 \\ 0 & 1 & \frac{9}{5} & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & \frac{8}{5} & -\frac{1}{5} & -\frac{7}{5} & 1 \end{array}\right]$$

Performing $\xrightarrow{\left(\frac{5}{8}\right)R_3}$, we get

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 5 & 1 & 0 & 0 \\ 0 & 1 & \frac{9}{5} & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & 1 & -\frac{1}{8} & -\frac{7}{8} & \frac{5}{8} \end{array}\right]$$

So we have three systems as follows :

$$\begin{bmatrix} 1 & 3 & 5 \\ 0 & 1 & \frac{9}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{2}{5} \\ -\frac{1}{8} \end{bmatrix} \quad \ldots (1)$$

$$\begin{bmatrix} 1 & 3 & 5 \\ 0 & 1 & \frac{9}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{5} \\ -\frac{7}{8} \end{bmatrix} \quad \ldots (2)$$

$$\begin{bmatrix} 1 & 3 & 5 \\ 0 & 1 & \frac{9}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \frac{5}{8} \end{bmatrix} \quad \ldots (3)$$

Writting down the corresponding system for (1), we get

$$x_1 + 3y_1 + 5z_1 = 1 \quad \ldots (i)$$

$$y_1 + \frac{9}{5} z_1 = \frac{2}{5} \quad \ldots (ii)$$

$$z_1 = -\frac{1}{8} \quad \ldots (iii)$$

Using $z_1 = -\frac{1}{8}$ in equation (ii) we get,

$$y_1 = \frac{2}{5} + \frac{9}{40} = \frac{5}{8}$$

Using $z_1 = -\frac{1}{8}$ and $y_1 = \frac{5}{8}$ in equation (i) we get,

$$x_1 = 1 - \frac{15}{8} + \frac{5}{8} = -\frac{1}{4}$$

Similarly writing the corresponding system for (2) and (3) and solving by back substitution method we obtain :

$$x_2 = \frac{1}{4},\ y_2 = \frac{11}{8},\ z_2 = -\frac{7}{8}$$

and $x_3 = \frac{1}{4},\ y_3 = -\frac{9}{8},\ z_3 = \frac{5}{8}$

Thus, $$A^{-1} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{5}{8} & \frac{11}{8} & -\frac{9}{8} \\ -\frac{1}{8} & -\frac{7}{8} & \frac{5}{8} \end{bmatrix}$$

Example 1.16 : Let $A = \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 3 & -5 & 4 \\ 6 & 7 & 5 & 11 \end{bmatrix}$

and the elementary matrix

$$E = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

obtained by performing the row operation $R_1 + (-2)\, R_2$ on I_3. Verify theorem (12)

Solution : We have

$$EA = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 3 & -5 & 4 \\ 6 & 7 & 5 & 11 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -7 & 12 & -8 \\ 0 & 3 & -5 & 4 \\ 6 & 7 & 5 & 11 \end{bmatrix}$$

If we perform the row operation $R_1 + (-2) R_2$ on A, we obtain a matrix

$$B = \begin{bmatrix} 1 & -7 & 12 & -8 \\ 0 & 3 & -5 & 4 \\ 6 & 7 & 5 & 11 \end{bmatrix}$$

Thus, $B = EA$

Example 1.17 : Let $A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 8 & 1 & 5 \\ 2 & -7 & -1 \\ 3 & 4 & 1 \end{bmatrix}$

(a) Find the elementary matrix E such that EA = B.

(b) Find the elementary matrix E_1 such that $E_1B = A$.

Solution : (a) If E results by row operation e on I_3, then EA = e(A), hence we need to find elementary row operation which when performed on A gives B. Clearly, this operation is R_{13}, the interchange of first row and third row of A.

Therefore, $$E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

(b) Here also, it is clear that A is obtained from B by performing the elementary row operation R_{13} on B. Therefore,

$$E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Example 1.18 : Consider the matrix

$$A = \begin{bmatrix} 2 & 0 \\ -3 & 1 \end{bmatrix}$$

(a) Find elementary matrices E_1 and E_2 such that $E_2E_1 A = I_2$

(b) Write A^{-1} as product of two elementary matrices.

Solution : (a) To find elementary matrices E_1 and E_2 such that $E_2E_1A = I$, we need to find two elementary row operations which when performed on A reduce it to I_2. We have

$$A = \begin{bmatrix} 2 & 0 \\ -3 & 1 \end{bmatrix} \xrightarrow{\left(\frac{1}{2}\right) R_1} \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \xrightarrow{R_2 + (3) R_1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then E_1 corresponds to the elementary row operation $\left(\frac{1}{2}\right) R_1$ and E_2 corresponds to the elementary row operation $R_2 + (3) R_1$. Therefore,

$$E_1 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \text{ and } E_2 = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$$

It is easy to see that

$$E_2 E_1 A = I_2 \qquad \ldots (*)$$

(b) From (*) of (a), we have,

$$(E_2 E_1) A = I_2$$

This shows that $E_2 E_1$ is inverse of A,

Therefore, $A^{-1} = E_2 E_1$

where, $E_1 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$ and $E_2 = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}$

Example 1.19 : Let $A = \begin{bmatrix} 1 & 2 \\ -3 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$

If e is elementary operation $e = R_{12}$, then verify e(AB) = e(A) B.

Solution : We have

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 11 & 17 \\ -7 & 29 & 45 \end{bmatrix}$$

Therefore, $e(AB) = \begin{bmatrix} -7 & 29 & 45 \\ 3 & 11 & 17 \end{bmatrix}$

Also $e(A)\,B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} -7 & 29 & 45 \\ 3 & 11 & 17 \end{bmatrix}$

Thus, $e(AB) = e(A)B$

Example 1.20 : Find the inverse of A by Gaussian elimination method, where

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

Solution : The augmented system is

$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 3 & 2 & 3 & 0 & 1 & 0 \\ 1 & 4 & 9 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_{13}} \left[\begin{array}{ccc|ccc} 1 & 4 & 9 & 0 & 0 & 1 \\ 3 & 2 & 3 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_2 + (-3)R_1}, \xrightarrow{R_3 + (-2)R_1} \left[\begin{array}{ccc|ccc} 1 & 4 & 9 & 0 & 0 & 1 \\ 0 & -10 & -24 & 0 & 1 & -3 \\ 0 & -7 & -17 & 1 & 0 & -2 \end{array}\right]$$

$$\xrightarrow{\left(-\frac{1}{10}\right)R_2} \left[\begin{array}{ccc|ccc} 1 & 4 & 9 & 0 & 0 & 1 \\ 0 & 1 & \frac{12}{5} & 0 & -\frac{1}{10} & \frac{3}{10} \\ 0 & -7 & -17 & 1 & 0 & -2 \end{array}\right]$$

$$\xrightarrow{R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 4 & 9 & 0 & 0 & 1 \\ 0 & 1 & \frac{12}{5} & 0 & -\frac{1}{10} & \frac{3}{10} \\ 0 & 0 & -\frac{1}{5} & 1 & -\frac{7}{10} & \frac{1}{10} \end{array}\right]$$

$$\xrightarrow{(-5)R_3} \left[\begin{array}{ccc|ccc} 1 & 4 & 9 & 0 & 0 & 1 \\ 0 & 1 & \frac{12}{5} & 0 & -\frac{1}{10} & \frac{3}{10} \\ 0 & 0 & 1 & -5 & \frac{7}{2} & -\frac{1}{2} \end{array}\right]$$

This is equivalent to the three systems

$$\begin{bmatrix} 1 & 4 & 9 \\ 0 & 1 & \frac{12}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 9 \\ 0 & 1 & \frac{12}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{10} \\ \frac{7}{2} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 9 \\ 0 & 1 & \frac{12}{5} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{3}{10} \\ -\frac{1}{2} \end{bmatrix}$$

Solving these three systems by back substitution method, we obtain

$$x_1 = -3, \quad y_1 = 12, \quad z_1 = -5$$
$$x_2 = \frac{5}{2}, \quad y_2 = -\frac{17}{2}, \quad z_2 = \frac{7}{2}$$
$$x_3 = -\frac{1}{2}, \quad y_3 = \frac{3}{2}, \quad z_3 = -\frac{1}{2}$$

Therefore,

$$A^{-1} = \begin{bmatrix} -3 & \frac{5}{2} & -\frac{1}{2} \\ 12 & -\frac{17}{2} & \frac{3}{2} \\ -5 & \frac{7}{2} & -\frac{1}{2} \end{bmatrix}$$

Exercise 1.2

1. If A is an $n \times n$ invertible matrix, then show that the system of linear equations $AX = B$ has a unique solution given by $X = A^{-1} B$.,
2. If A is an $n \times n$ invertible matrix, then show that the system $AX = O$ has only trivial solution.
3. If A is an $n \times n$ matrix and the homogeneous system $AX = O$ has only trivial solution, then prove that A must be row equivalent to I_n.
4. Let the system of linear equations $AX = B$ have a solution X_o. Then prove that the general solution of the system is given by $X = X_o + X'$, where X' is any solution of the homogeneous system $AX = O$.
5. If A and B are $m \times n$ and $n \times p$ matrices respectively, and if e is an elementary row operation and e' is an elementary column operation, then prove that
   (a) e performed on the product AB is the same as e performed on the prefactor A of AB.
   (b) e' performed on the product AB is the same as e' performed on the post factor B of AB.
6. If the elementary matrix E results from performing a certain row operation e on I_m and if A is an $m \times n$ matrix, then prove that the product EA is the matrix that results when this same row operation is performed on A.
7. Prove that every elementary matrix is invertible and the inverse is also an elementary matrix.
8. If an $n \times n$ matrix A is row equivalent to I_n, then prove that A is invertible.
9. Prove that every invertible matrix is a product of elementary matrices.

10. Prove that a square matrix is invertible if and only if its row echelon form is invertible.
11. For the following matrices find the row echelon form and reduced row echelon form.

(a) $\begin{bmatrix} 2 & 0 & 2 & 1 \\ 3 & -1 & 4 & 7 \\ 6 & 1 & -1 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & -1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -2 & -2 \\ -1 & 2 & -4 & 1 & 1 \\ 3 & 0 & 0 & -3 & -3 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -4 \\ -4 & 1 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 2 & 3 & 1 & 0 & 4 \\ 3 & 1 & 2 & -1 & 1 \\ 4 & -1 & 3 & -2 & -2 \\ 5 & 4 & 3 & -1 & 6 \end{bmatrix}$

12. Solve the following systems of linear equations by :
    (a) Gaussian elimination method and
    (b) Gauss-Jordan elimination method

(i)
$$\begin{aligned} 3x_1 + 4x_2 - x_3 &= -1 \\ x_1 + 3x_3 &= 2 \\ 2x_1 + 5x_2 - 4x_3 &= 3 \end{aligned}$$

(ii)
$$\begin{aligned} x + y + z &= 5 \\ x + y - 4z &= 10 \\ -4x + y + z &= 0 \end{aligned}$$

(iii)
$$\begin{aligned} x_1 + x_2 + 2x_3 &= 9 \\ 2x_1 + 4x_2 - 3x_3 &= 1 \\ 3x_1 + 6x_2 - 5x_3 &= 0 \end{aligned}$$

(iv)
$$\begin{aligned} 2x_1 + 3x_2 &= -2 \\ 2x_1 + x_2 &= 1 \\ 3x_1 + 2x_2 &= 1 \end{aligned}$$

(v)
$$\begin{aligned} -2x_3 + 7x_5 &= 12 \\ 2x_1 + 4x_2 - 10x_3 + 6x_4 + 12x_5 &= 28 \\ 2x_1 - 4x_2 - 5x_3 + 6x_4 + 5x_5 &= -1 \end{aligned}$$

(vi)
$$\begin{aligned} x_1 - x_2 + 2x_3 - x_4 &= -1 \\ 2x_1 + x_2 - 2x_3 - 2x_4 &= -2 \\ -x_1 + 2x_2 - 4x_3 + x_4 &= 1 \\ 3x_1 + (-3)\, x_4 &= -3 \end{aligned}$$

(vii)
$$\begin{aligned} 2x_1 + 4x_2 + x_3 + 7x_4 &= 7 \\ x_3 + 2x_4 - x_5 &= 4 \\ x_4 - x_5 &= 3 \\ x_3 + 3x_4 - 2x_5 &= 7 \end{aligned}$$

(viii)
$$\begin{aligned} 2x_1 + 2x_2 - x_3 + x_5 &= 0 \\ -x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 - x_5 &= 0 \\ x_3 + x_4 + x_5 &= 0 \end{aligned}$$

(ix)
$$\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ x_1 + 2x_2 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

(x)
$$\begin{aligned} 2x_2 + 2x_3 + 4x_4 &= 0 \\ x_1 - x_3 - 3x_4 &= 0 \\ 2x_1 + 3x_2 + x_3 + x_4 &= 0 \\ -2x_1 + x_2 + 3x_3 - 2x_4 &= 0 \end{aligned}$$

(xi)
$$\begin{aligned} 2x - y - 3z &= 0 \\ -x + 2y - 3z &= 0 \\ x + y + 4z &= 0 \end{aligned}$$

13. Show that the system of equations
$$\begin{aligned} x + y + 2z &= a \\ x + z &= b \\ 2x + y + 3z &= c \end{aligned}$$
is consistent only if $a + b = c$.

14. For which values of 'a' does the following systems of linear equations have non-trivial solution ?
$$\begin{aligned} (a-3)\, x + y &= 0 \\ x + (a-3)\, y &= 0 \end{aligned}$$

15. In each of the following find the condition that b's must satisfy for the system to be consistent.

(i)
$$\begin{aligned} x_1 - 2x_2 + 5x_3 &= b_1 \\ 4x_1 - 5x_2 + 8x_3 &= b_2 \\ -3x_1 + 3x_2 - 3x_3 &= b_3 \end{aligned}$$

(ii)
$$\begin{aligned} x_1 - 2x_2 - x_3 &= b_1 \\ -4x_1 + 5x_2 + 2x_3 &= b_2 \\ -4x_1 + 7x_2 + 4x_3 &= b_3 \end{aligned}$$

(iii)
$$\begin{aligned} x_1 - x_2 + 3x_3 + 2x_4 &= b_1 \\ -2x_1 + x_2 + 5x_3 + x_4 &= b_2 \\ -3x_1 + 2x_2 + 2x_3 - x_4 &= b_3 \\ 4x_1 - 3x_2 + x_3 + 3x_4 &= b_4 \end{aligned}$$

(iv)
$$\begin{aligned} x_1 + x_2 + 2x_3 &= b_1 \\ x_1 + x_3 &= b_2 \\ 2x_1 + x_2 + 3x_3 &= b_3 \end{aligned}$$

(v)
$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= b_1 \\ 2x_1 - 15x_2 + 3x_3 &= b_2 \\ x_1 + 8x_3 &= b_3 \end{aligned}$$

16. For which values of 'a' does the following system have (i) Unique solution ? (ii) Infinitely many solutions ?
$$\begin{aligned} x_1 + x_2 + x_3 &= 4 \\ x_3 &= 2 \\ (a^2 - 4)\, x_3 &= a - 2 \end{aligned}$$

17. Which of the following are elementary matrices ?

(i) $\begin{bmatrix} 1 & \sqrt{5} \\ 0 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 1 \\ 0 & -5 \end{bmatrix}$ (iii) $\begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix}$ (iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (vi) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -19 \\ 0 & 0 & 1 \end{bmatrix}$

18. Let $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & 1 \\ 4 & 5 & -1 \end{bmatrix}$

and let $E_1 = \begin{bmatrix} 1 & -5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Identify the elementary row operations that produce E_1, E_2 and E_3 and verify theorem (12).

19. Consider the matrices

$$A = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 8 & 1 & 5 \end{bmatrix}, B = \begin{bmatrix} 8 & 1 & 5 \\ 2 & -7 & -1 \\ 3 & 4 & 1 \end{bmatrix}, C = \begin{bmatrix} 3 & 4 & 1 \\ 2 & -7 & -1 \\ 2 & -7 & 3 \end{bmatrix}$$

Find elementary matrices E_1, E_2, E_3 and E_4 such that (a) $E_1A = B$, (b) $E_2 B = A$, (c) $E_3A = C$, (d) $E_4C = A$.

20. In each part perform the stated row operation on

$$\begin{bmatrix} 2 & -1 & 0 \\ 4 & 5 & -3 \\ 1 & -4 & 7 \end{bmatrix}$$

by multiplying A on the left by a suitable elementary matrix. Check your answer in each case by performing the row operation directly on A.

(a) Interchange of first and third row.

(b) Multiply the third row by $-\frac{2}{3}$.

(c) Add twice the first row to the second row.

21. Express the matrix

$$A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix}$$

in the form $A = E_1 E_2 E_3 R$, where E_1, E_2 and E_3 are elementary matrices and R is row echelon form.

22. If $A = \begin{bmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

is an elementary matrix, then show that atleast one of the entry in first row must be zero.

23. Compute the inverse of the matrix

$$A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 1 & 1 \\ 1 & 3 & 5 \end{bmatrix}$$

by Gaussian elimination method and use the result to solve the system of equations:

$$\begin{aligned} 3x + 2y + 4z &= 7 \\ 2x + y + z &= 7 \\ x + 3y + 5z &= 2 \end{aligned}$$

24. Compute the inverse of the matrix

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & 4 \\ 1 & 2 & 2 \end{bmatrix}$$

by Gaussian elimintaion method.

25. Find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -4 \\ -4 & 1 & 1 \end{bmatrix}$$

by Gaussian elimination method and use the result to solve the system of equations:

$$\begin{aligned} x + y + z &= 5 \\ x + y - 4z &= 10 \\ -4x + y + z &= 0 \end{aligned}$$

Answers (1.2)

11. (a) $\begin{bmatrix} 1 & 0 & 1 & \frac{1}{2} \\ 0 & 1 & -1 & -\frac{11}{2} \\ 0 & 0 & 1 & -\frac{5}{12} \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & \frac{11}{12} \\ 0 & 1 & 0 & -\frac{71}{12} \\ 0 & 0 & 1 & -\frac{5}{12} \end{bmatrix}$

(b) $\begin{bmatrix} 1 & -1 & 2 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & -2 & -1 & -1 & -3 \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} & \frac{10}{7} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -\frac{9}{7} & -\frac{3}{7} & 0 \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

12. (i) $x_1 = -\frac{73}{10}$, $x_2 = 6$, $x_3 = \frac{31}{10}$.
    (ii) $x = 1, y = 5, z = -1$.
    (iii) $x_1 = 1, x_2 = 2, x_3 = 3$.
    (iv) The system is inconsistent.
    (v) $x_1 = 7 - 2s - 3t$, $x_2 = s$, $x_3 = 1$, $x_4 = t$, $x_5 = 2$; $s, t \in \mathbb{R}$.
    (vi) $x_1 = 1 + s$, $x_2 = 2r$, $x_3 = r$, $x_4 = s$; $s, r \in \mathbb{R}s$
    (vii) $x_1 = -6 - 2s - 3t$, $x_2 = s$, $x_3 = -2 - t$, $x_4 = 3 + t$, $x_5 = t$; $s, t \in \mathbb{R}$
    (viii) $x_1 = -s - t$, $x_2 = s$, $x_3 = t$, $x_4 = 0$, $x_5 = t$; $s, t \in \mathbb{R}$
    (ix) $x_1 = 0 = x_2 = x_3$, only trivial solution
    (x) $x_1 = t$, $x_2 = -t$, $x_3 = t$, $x_4 = 0$; $t \in \mathbb{R}$.
    (xi) $x = 0 = y = z$ only trivial solution
14. $a = 4$, $a = 2$
15. (i) $b_1 = b_2 + b_3$ (ii) No restriction on b's.
    (iii) $b_1 = b_3 + b_4$ and $b_2 = 2b_3 + b_4$ (iv) $b_3 = b_1 + b_2$
    (v) No restriction.
16. Infinitely many solutions if $a = 2$ or $a = -\frac{3}{2}$; none otherwise.
17. (i), (iii) and (vi) are elementary matrices
    (ii), (iv) and (v) are not elementary matrices
19. $E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$,
    $E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$
21. $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $E_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
    $R = \begin{bmatrix} 1 & 3 & 3 & 8 \\ 0 & 1 & 7 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

23. $A^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ -\frac{9}{8} & \frac{11}{8} & \frac{5}{8} \\ \frac{5}{8} & -\frac{7}{8} & -\frac{1}{8} \end{bmatrix}$ $x = \frac{9}{4}$, $y = -\frac{9}{8}$, $z = \frac{5}{8}$

24. $A^{-1} = \begin{bmatrix} \frac{6}{5} & -\frac{2}{5} & \frac{1}{5} \\ -\frac{1}{5} & -\frac{1}{10} & \frac{3}{10} \\ -\frac{2}{5} & \frac{3}{10} & \frac{1}{10} \end{bmatrix}$

25. $A^{-1} = \begin{bmatrix} \frac{1}{5} & 0 & -\frac{1}{5} \\ \frac{3}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & -\frac{1}{5} & 0 \end{bmatrix}$, $x = 1, y = 5, y = -1.$

1.8 LU-Factorization

The Gaussian-elimination and Gauss-Jordan methods are very useful to solve the linear systems LU-factorization method is the most widely used method on computers. The main reason for the popularity of this method is that it provides the cheapest way of solving systems of linear equations for which we repeatedly have to change the right side. This type of situation occurs often in applied problems. For example, an electric utility company must determine the inputs (the unknowns) needed to produce some required outputs (the right sides). The inputs and outputs must be related by linear equations, whose coefficient matrix is fixed, while the right side changes from day-to-day, or even hour-to-hour. In process of LU-factorization, we **decompose** the **coefficient** matrix as product of a lower triangular matrix and an upper triangular matrix.

LU-Factorization or LU-Decomposition :

We first describe the procedure of LU-factorization, then we will describe how to obtain the LU-factorization of a given matrix.

Let A be an $n \times n$ matrix and suppose L is lower triangular matrix and U is an upper triangular matrix such that

$$A = LU$$

In this case we say that A has an LU-factorization or an LU-decomposition. Let us see the LU-factorization of a matrix can be used efficiently to solve the system of linear equations.

$$AX = B$$

Substituting LU for A, we have,

$$(LU)\,X = B$$

by associativity of matrix multiplication, this can be written as

$$L(UX) = B$$

Taking UX = Z, this matrix becomes

$$LZ = B$$

Since L is in lower triangular form, we solve directly for Z by forward substitution. Once Z is determined, and since U is upper triangular matrix, we solve UX = Z by back substitution. In short, if an $n \times n$ matrix A has an LU-factorization, then the solution of AX = B can be determined by a forward substitution followed by back substitution.

Now we describe how to obtain LU-factorization of a given $n \times n$ matrix, we illustrate it step-by-step by an example in the following.

Let, $$A = \begin{bmatrix} -5 & 4 & 0 & 1 \\ -30 & 27 & 2 & 7 \\ 5 & 2 & 0 & 2 \\ 10 & 1 & -2 & 1 \end{bmatrix}$$

Procedure **Matrices used**

Step 1 : To bring zeros below the first diagonal entry of A, add (–6) times the first row of A to the second row of A; add 1 times the first row of A to the third row of A, and 2 times the first row of A to the fourth row of A. We call new resulting matrix as U_1.

$$U_1 = \begin{bmatrix} -5 & 4 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 6 & 0 & 3 \\ 0 & 9 & -2 & 3 \end{bmatrix}$$

We construct simultaneously a lower triangular matrix L_1 with its 1's on main diagonal to record the row operations. Enter the negatives of the multiples used in the row operations in the first column of L_1, below the first diagonal entry of L_1.

$$L_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 \\ -1 & * & 1 & 0 \\ -2 & * & * & 1 \end{bmatrix}$$

Step 2 : To bring zeros below the second diagonal entry of U_1, add (–2) times the second row of U_1 to the third row of U_1, (–3) times the second row of U_1 to the fourth row of U_1. Call new matrix as U_2.

$$U_2 = \begin{bmatrix} -5 & 4 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & -4 & 1 \\ 0 & 0 & -8 & 0 \end{bmatrix}$$

Enter the negatives of the multipliers from the row operations below the second diagonal entry of L_1, call the new matrix as L_2.

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ -2 & 3 & * & 1 \end{bmatrix}$$

Step 3 : To bring zeros below the third diagonal entry of U_2, add (–2) times the third row of U_2 to the fourth row of U_2 and call the new matrix as U_3.

$$U_3 = \begin{bmatrix} -5 & 4 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

Enter the negative of the multiplier below the third diagonal entry of L_2. Call the new matrix as L_3.

$$L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ -2 & 3 & 2 & 1 \end{bmatrix}$$

Let $L = L_3$ and $U = U_3$. Then the product LU gives the original matrix A (students are requested to verify this).

Note : If the diagonal entry $a_{11} = 0$ in A, then the above procedure fails or second diagonal entry in U_1 is zero then again this process fails. In such cases we have to use alternating the sequence of given linear equations and use this method.

Example 1.21 : Consider the linear system :

$$\begin{aligned} -5x_1 + 4x_2 + 0\cdot x_3 + x_4 &= -17 \\ -30x_1 + 27x_2 + 2x_3 + 7x_4 &= -102 \\ 5x_1 + 2x_2 + 0\cdot x_3 + 2x_4 &= -7 \\ 10x_1 + x_2 - 2x_3 + x_4 &= -6 \end{aligned}$$

Solution : Here the matrix of coefficient is

$$A = \begin{bmatrix} -5 & 4 & 0 & 1 \\ -30 & 27 & 2 & 7 \\ 5 & 2 & 0 & 2 \\ 10 & 1 & -2 & 1 \end{bmatrix}$$

This matrix A has LU-factorization, where

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ -2 & 3 & 2 & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} -5 & 4 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

We have, $$B = \begin{bmatrix} -17 \\ -102 \\ -7 \\ -6 \end{bmatrix}$$

To solve the system using this LU-factorization, we proceed as AX = B, then (LU) X = B. First let UX = Z and let us solve LZ = B.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 6 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ -2 & 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} -17 \\ -102 \\ -7 \\ -6 \end{bmatrix}$$

By forward substitution, we obtain,

$$z_1 = -17,\ 6z_1 + z_2 = -102$$

$$-z_1 + 2z_2 + z_3 = -7$$

$$-2z_1 + 3z_2 + 2z_3 + z_4 = -6$$

$$\therefore \quad z_1 = -17$$

$$z_2 = -102 - 6z_1 = -102 - 6 \times (-17) = 0$$

$$z_3 = -7 + z_1 - 2z_2 = -7 - 17 - 0 = -24$$

$$z_4 = -6 + 2z_1 - 3z_2 - 2z_3 = -6 - 34 - 0 + 48 = 8$$

Next we solve, $UX = Z$

$$\begin{bmatrix} -5 & 4 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -17 \\ 0 \\ -24 \\ 8 \end{bmatrix}$$

By backward substitution, we obtain,

$$x_4 = \frac{8}{-2} = -4$$

$$x_3 = \frac{-24 - x_4}{-4} = \frac{-24 + 4}{-4} = 5$$

$$x_2 = \frac{-2x_3 - x_4}{3} = \frac{-10 + 4}{3} = -2$$

$$x_1 = \frac{-17 - 4x_2 - x_4}{-5} = \frac{-17 + 8 + 4}{-5} = 1$$

Thus, $x_1 = 1$, $x_2 = -2$, $x_3 = 5$ and $x_4 = -4$ is the solution of the given system of linear equations obtained by LU-factorization method.

Exercise 1.3

In the following exercises 1 to 4, let AX = B be a linear system. In each case, LU-factorization of the coefficient matrix is given, solve the system using forward substitution followed by a back substitution.

1. $A = \begin{bmatrix} 2 & 8 & 0 \\ 2 & 2 & -3 \\ 1 & 2 & 7 \end{bmatrix}, B = \begin{bmatrix} 18 \\ 3 \\ 12 \end{bmatrix}, L = \begin{bmatrix} 2 & 0 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & 4 \end{bmatrix}, U = \begin{bmatrix} 1 & 4 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

2. $A = \begin{bmatrix} 6 & -2 & -4 & 4 \\ 3 & -3 & -6 & 1 \\ -12 & 8 & 21 & -8 \\ -6 & 0 & -10 & 7 \end{bmatrix}, B = \begin{bmatrix} 2 \\ -4 \\ 8 \\ -43 \end{bmatrix}$

$L = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ -4 & 2 & 1 & 0 \\ -2 & -1 & -2 & 2 \end{bmatrix}$ and $U = \begin{bmatrix} 3 & -1 & -2 & 2 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 5 & -2 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

3. $A = \begin{bmatrix} 8 & 12 & -4 \\ 6 & 5 & 7 \\ 2 & 1 & 6 \end{bmatrix}, B = \begin{bmatrix} -36 \\ 11 \\ 16 \end{bmatrix}, L = \begin{bmatrix} 4 & 0 & 0 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & 3 & -1 \\ 0 & -2 & 5 \\ 0 & 0 & 2 \end{bmatrix}$

4. $A = \begin{bmatrix} 4 & 2 & 1 & 0 \\ -4 & -6 & 1 & 3 \\ 8 & 16 & -3 & -4 \\ 20 & 10 & 4 & -3 \end{bmatrix}, B = \begin{bmatrix} 6 \\ 13 \\ -20 \\ 15 \end{bmatrix},$

$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -3 & 1 & 0 \\ 5 & 0 & -1 & 1 \end{bmatrix}, U = \begin{bmatrix} 4 & 2 & 1 & 0 \\ 0 & -4 & 2 & 3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

In the following exercises 5 to 7, find an LU-factorization of the coefficient matrix of the system of linear equations AX = B. Solve the system using a forward substitution followed by a back substitution.

5. $A = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 5 & 10 \\ 4 & 8 & 2 \end{bmatrix}, B = \begin{bmatrix} 6 \\ 16 \\ 2 \end{bmatrix}$

6. $A = \begin{bmatrix} 4 & 2 & 3 \\ 2 & 0 & 5 \\ 1 & 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$

7. $A = \begin{bmatrix} 2 & 3 & 0 & 1 \\ 4 & 5 & 3 & 3 \\ -2 & -6 & 7 & 7 \\ 8 & 9 & 5 & 21 \end{bmatrix}$,

$B = \begin{bmatrix} -2 \\ -2 \\ -16 \\ -66 \end{bmatrix}$

Answers (1.3)

1. $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ 2. $\begin{bmatrix} 9/2 \\ 69/10 \\ -6/5 \\ -4 \end{bmatrix}$

3. $\begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}$ 4. $\begin{bmatrix} 2 \\ -1 \\ 0 \\ 5 \end{bmatrix}$

5. $L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -2 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & 3 & 4 \\ 0 & -1 & 2 \\ 0 & 0 & -2 \end{bmatrix}, X = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$

6. $L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 1/4 & -3/2 & 1 \end{bmatrix}, U = \begin{bmatrix} 4 & 2 & 3 \\ 0 & -1 & 7/2 \\ 0 & 0 & 11/2 \end{bmatrix}, X = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$

7. $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 3 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}, U = \begin{bmatrix} 2 & 3 & 0 & 1 \\ 0 & -1 & 3 & 1 \\ 0 & 0 & -2 & 5 \\ 0 & 0 & 0 & 4 \end{bmatrix}, X = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -4 \end{bmatrix}$

Miscellaneous Exercise

(A) Theory Questions :

1. Define the system of linear equation.
2. Define : (i) Matrices, (ii) Scalar multiplication of matrices, (iii) Zero matrix, (iv) Trace of matrix, (v) Elementary matrix.
3. State and explain types of matrices with one example each.
4. State the properties of adjoint of a matrix.
5. Explain or write short note on LU-factorization.

(B) Numerical Problems :

1. Solve the following system by using Gauss-Jordan elimination method.

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 1 \\ 3x_1 + 4x_4 &= 1 \\ x_1 - 4x_2 - 2x_3 - 2x_4 &= 0 \end{aligned}$$

2. For which values of k will the following system have (i) No solution ? (ii) Exactly one solution ? (iii) Infinitely many solutions ?

$$\begin{aligned} 4x + y + (k^2 - 14)\, z &= k + 2 \\ x + 2y - 3z &= 4 \\ 3x - y + 5z &= 2 \end{aligned}$$

[**Ans.** (i) $k = -4$, (ii) $k \neq \pm 4$, (iii) $k = 4$]

3. Compute the inverse of A by Gaussian elimination method, where

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 2 & 4 \end{bmatrix} \quad \left[\textbf{Ans. } A^{-1} = \begin{bmatrix} -1/4 & 1/4 & 1/4 \\ 5/8 & 11/8 & -9/8 \\ -1/8 & -7/8 & 5/8 \end{bmatrix} \right]$$

4. Find the general solution of

$$\begin{aligned} 3x_1 + 4x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$$

[**Ans.** $x_1 = -3t,\ x_2 = 2t,\ x_3 = t,\ t \in \mathbb{R}$]

5. If $A = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}$, show that $A^2 = A$ i.e. A is idempotent.

6. Solve the system of linear equations

$$\begin{aligned} x + y + z &= 1 \\ 2x - y + z &= 2 \end{aligned}$$

$\left[\textbf{Ans. } x = 1 - \frac{2}{3}t,\ y = -\frac{1}{3}t,\ z = t,\ t \in \mathbb{R} \right]$

7. Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & b & -1 & 0 \\ c & d & 0 & -1 \end{bmatrix}$. Using given partition of A, prove that $A^2 = I$.

8. Show that +ve odd integral powers of a skew symmetric matrix is skew symmetric while +ve even integral power is symmetric matrix.

9. If X and Y are two symmetric matrices, show that XYX is also symmetric. Is XY symmetric matrix always ? Justify your answer.

10. Show that (i) (AB) C = A (BC), (ii) A (B + C) = AB + AC, where A, B, C are matrices with multiplication compability.

11. Show that : $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & - & -1 \end{bmatrix}$ are inverses of each other.

12. Suppose A and B are orthogonal matrices show that A^T, A^{-1}, AB are also orthogonal.

13. Solve each of the following systems :

    (a) $2x - 5y = 11$

    $2x + 4y = 5$

    Ans. $x = 3, y = -1$

    (b) $2x - 3y = 8$

    $-6x + 9y = 6$

    Ans. No solution.

    (c) $2x - 3y = 8$

    $-4x + 6y = -16$

    Ans. $x = \frac{3}{2} + 4, \quad y = t$

14. Solve each of the following system :

    (a) $x + 2y - 4z = -4$

    $2x + 5y - 9z = -10$

    $3x - 2y + 3z = 11$

    Ans. $x = 2, y = -1, z = 1$

(b) $x + 2y - 3z = -1$

$-3x + y - 2z = -7$

$5x + 3y - 4z = 2$

Ans. No solution.

(c) $x + 2y - 3z = 1$

$2x + 5y - 8z = 4$

$3x + 8y - 13z = 7$

Ans. $x = -3 - a,\ y = 2 + 2a,\ z = a \Rightarrow$ a is scalar.

❑❑❑

Chapter 2...

Generalised Inverse

Synopsis

2.1 Introduction
2.2 Generalized Inverse : Definition
2.3 Generalized inverse of Arbitrary matrices
2.4 Applications of generalized inverse in statistics
2.5 Moore-Penrose generalized inverse
2.6 Existence of generalized inverse by Different Methods
2.7 Singular Value Decomposition (SVD)
2.8 SVD Algorithm
2.9 Full Rank Factorization Method
Solved Examples

Objectives

At the end of the this chapter.

(1) You will be able to understand, existence of inverse of a square matrix if given matrix is singular.

(2) You will be able to understand types of generalised inverse.

(3) You will be able to understand what is singular value decomposition.

(4) You will be able to find full rank matrices using factorization method.

David S. Moore

David Sheldon Moore is an American statistician, who is known for his leadership of statistics education for many decades. David S. Moore received his A.B. from Princeton University and the Ph.D. from Cornell University in mathematics. In statistics education, David S. Moore is the author of a series of influential textbooks in Statistical Science, which use only high school algebra: Introduction to the Practice of Statistics (with George McCabe), of An Introduction to the Basic Practice of Statistics, and of Statistics: Concepts and Controversies. In statistical science, David S. Moore has done research in the asymptotic theory of robust and non-parametric statistics. Professor Moore was the 1998 President of the American Statistical Association. David S. Moore is a retired in 2004.

Roger Penrose

Sir Roger Penrose OM FRS (born 8 August 1931) is an English mathematical physicist, mathematician and philosopher of science. He is the Emeritus Rouse Ball Professor of Mathematics at the Mathematical Institute of the University of Oxford, as well as an Emeritus Fellow of Wadham College. Penrose is known for his work in mathematical physics, in particular for his contributions to general relativity and cosmology. He has received several prizes and awards, including the 1988 Wolf Prize for physics, which he shared with Stephen Hawking for their contribution to our understanding of the universe.

2.1 Introduction

Let A be a general $m \times n$ matrix then a natural question is when we can solve

$$Ax = y \quad \text{for } x \in R^n, \text{ given } y \in R^m \qquad \ldots (2.1)$$

If A is a square matrix ($m = n$) and A has an inverse, then (2.1) holds if and only if $x = A^{-1}y$. However, A may be $m \times n$ with $m \neq n$ or A may be a square matrix that is not invertible.

If A is not invertible, then equation (2.1) may have no solutions (that is, y may not be in the range of A) and if there are solutions, then there may be many different solutions.

For example, assume $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Then $A \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, so that A is not invertible. It would be useful to have a characterization of those $y \in R^2$ for which it is possible to find a solution of $Ax = y$ and if $Ax = y$ is a solution, to find all possible solutions. It is easy to answer these questions directly for a 2×2 matrix, but not if A were 5×3 or 10×3.

The answers to questions can be found in general from the notion of a *generalized inverse* of a matrix.

2.2 Generalized Inverse : Definition

If A is an $m \times n$ matrix, then G is a *generalized inverse* of A is an $n \times m$ matrix G with

$$AGA = A \qquad \ldots (2.2)$$

Note : A generalised inverse is denoted by A^- i.e. $G = A^-$.

(1) If A has an inverse in the usual sense, that is if A is $n \times n$ and has a two-sided inverse (A^{-1}), then pre and post multiplying by A^{-1} on both sides of (2.2) we get

$$A^{-1}(AGA)A^{-1} = A^{-1}AA^{-1}$$
$$(A^{-1}A)G(AA^{-1}) = A^{-1}(AA^{-1})$$
$$I_nGI_n = A^{-1}I_n \text{ is } G = A^{-1}.$$

Thus, if A^{-1} exists, then $G = A^{-1}$. This justifies the term *generalized* inverse.

(2) For any $m \times n$ matrix A has at least one generalized inverse G however A is $n \times n$ and A is invertible, there are many different generalized inverses G, so that G is not unique.

(3) Using (2.2) we can see that AGAG = AG and GAGA = GA. In general, a square matrix P that satisfies $P^2 = P$ is called a *projection matrix*. Thus both AG and GA are projection matrices. Since A is $m \times n$ and G is $n \times m$, AG is an $m \times m$ projection matrix and GA is $n \times n$ projection matrix.

(4) In general, if P is a projection matrix, then $P = P^2$ implies Py = P(Py) and Pz = z for all z = Py in the range of P. That is, if P is $n \times n$, P moves any $x \in R^n$ into $V = \{Px : x \in R^n\}$ (the range of P) and then keeps it at the same place.

If $x \in R^n$, then y = Px and z = x − Px = (1 − P) x satisfies x = y + z, Py = y and $Pz = P(x - Px) = Px - P^2x = 0$. Since then Px = P(y + z) = y, we can say that P projects R^n onto its range V along the space $W = \{x : Px = 0\}$.

The two projections AG and GA both appear in the next result, which shows how generalized inverse can be used to solve matrix equations.

Theorem 1 : Let A by an $m \times n$ matrix and assume that G is a generalized inverse of A then, for any fixed $y \in R^n$,

(i) the equation

$$Ax = y, \quad x \in R^n \qquad \ldots (2.3)$$

has a solution $x \in R^n$ if and only if AGy = y (that is, if and only if y is in the range of the projection (AG).

(ii) If Ax = y has solutions, then x is a solution of Ax = y if and only if

$$x = Gy + (I - GA) z \quad \text{for } z \in R^n \qquad \ldots (2.4)$$

Remark : If we want particular solution of Ax = y for y in the range of A, we can take x = Gy.

Proof : If y is in the range of the projection AG, that is if (AG) y = y, then A(Gy) = y and x = Gy is a solution of Ax = y. Conversely, if Ax = y, then GAx = Gy and AGAx = AGy = (AG)y, while AGA = A implies AGAx = Ax = y. Thus (AG) y = y. Thus, if Ax = y has any solutions for a given $y \in R^m$, then x = Gy is a particular solution.

Proof of part (ii) : This has two parts : First, if AGy = y, then all of the vectors in (2.4) are solutions of Ax = y. Second, that (2.4) contains all possible solutions of Ax = y.

If $AGy = y \;\&\; x = Gy + (I - GA) z$, then Ax = AGy + A(I − GA) z = y + (A − AGA) z = y, so that any $x \in R^n$ that satisfies (2.4) with AGy = y is a solution of Ax = y.

Conversely, if Ax = y, let z = x. Then the right-hand side of (2.4) is Gy + (I − GA) x = Gy + x − G(Ax) = Gy + x − Gy = x, so that any solution x of Ax = y is given by (2.4) with z = x.

Note : The generalized inverse is also referred to as the "conditional inverse", "Pseudo inverse" and 'g-inverse".

Example : Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Set $G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Then

$$AGA = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} = A$$

so that G is a generalized inverse of A. Then two projections appearing in Theorem 1 are

$$AG = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} \quad \text{and} \quad GA = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

In this case

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x + 6y \end{pmatrix} = (x + 2y) \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Thus $Ax = y$ has a solution x only if $y = c \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. On the other hand,

$$AG \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 3x \end{pmatrix} = x \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

so that the range of the projection AG is exactly the set of vectors $\left\{ c \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$.

The theorem then says that if $y = c \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, then the set of solutions of $Ax = y$ is exactly.

$$\begin{aligned} x &= Gc \begin{pmatrix} 1 \\ 3 \end{pmatrix} + (I - GA) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= c \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \right) - \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= c \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = c \begin{pmatrix} 1 \\ 0 \end{pmatrix} + z_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} \qquad \ldots (2.5) \end{aligned}$$

It is easy to check that $Ax = y = c \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ for all x in (2.5).

Theorem 2 : A generalized inverse always exists although it is not unique in general.

Proof : Assume rank (A) = r. According to the singular-value decomposition we have,

$$A_{m \times n} = U_{m \times r} \cdot L_{r \times r} \cdot V'_{r \times n}$$

with $U'U = I_r$ and $V'V = I_r$ and

$$L = \text{diag}\,(l_1, \ldots, l_r), \quad l_i > 0.$$

Then $$A^{-1} = V \begin{pmatrix} L^{-1} & X \\ Y & .Z \end{pmatrix} U'$$

(X, Y and Z are arbitrary matrices of suitable dimensions) is a g-inverse of A. Using Theorem, namely,

$$A = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$$

with X non-singular, we have

$$A^{-1} = \begin{pmatrix} X^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

as a special g-inverse.

2.3 Generalized Inverse of Arbitrary Matrices

Let A be an arbitrary m × n matrix; that is, with n columns and m rows. Then we can write

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (v_1 \quad v_2 \quad \ldots \quad v_n) = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \cdots \\ \cdots \\ \omega_m \end{pmatrix}$$

where v_i are the columns of A and w_j are the rows.

In general, the *column rank* of A (call it r_c) is the dimension of the vector space in R^m that is spanned by the columns $\{v_i\}$ of A, and the *row rank* of A (call it r_r) is the dimension of the vector space in R^n that is spanned by the rows $\{w_j\}$ of A. That is, r_c is the largest number of linearly-independent columns v_i in R^m and r_r is the largest number of linearly-independent rows w_j in R^n. Then $r_r \le m$, since the largest number of linearly independent vectors in R^m is m and $r_c \le n$ since there are only n columns to begin with. Thus $r_c \le \min\{m, n\}$. By the same arguments $r_r \le \min\{m, n\}$.

It can be shown that $r_c = r_r$ for any m × n matrix, so that the row rank and the column rank of an arbitrary matrix A are the same. The common value $r = r_c = r_r \le \min\{m, n\}$ is called the rank of A.

Let A be an m × n matrix with rank (A) = $r \le \min\{m, n\}$ then, we can show that, after a suitable rearrangement of rows and columns, A can be written in partitioned form as

$$A = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} \qquad \ldots (2.6)$$

where, X is r × r and invertible, Y is r × (n − r), Z is (m − r) × r and W is (m − r) × (n − r). In fact, we can prove the following presentation theorem for general matrices :

Theorem 3 : Let A is an m × n matrix with rank r then the rows and columns can be permuted so that it can be written in the partitioned form where X is r × r and invertible. In that case $W = ZX^{-1}Y$, so that

$$A = \begin{bmatrix} X & Y \\ Z & ZX^{-1}Y \end{bmatrix} \qquad \ldots (2.7)$$

(Note that X, Y, Z, W in (2.6) and (2.7) are *matrices*, not numbers. Some of the entries Y, Z, W in (2.6) may be empty, in this case they do not appear, for example if m = n and A is invertible.)

Remarks : If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix of numbers with $a > 0$ but $r = \text{rank}(A) = 1$, then $\det(A) = ad - bc = 0$. This implies $d = bc/a$. We cannot write bc/a for matrices, but (2.7) with $d = ba^{-1}c$ is the appropriate generalization for matrices. The matrix $d = ca^{-1}b$ and b is $r \times (n - r)$.

Example : Let $A = xy'$ be the product of vectors $x \in R^m$ and $y \in R^n$, so that A is $m \times n$. Assume $x_1 \neq 0$ and $y_1 \neq 0$. Then rank $(A) = 1$ since every row of A is a multiple of y and every column of A is a multiple of x. In this case, we can write

i.e.
$$A = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} (y_1 \quad y_2 \ \ldots \ y_n)$$

$$A = \left[\begin{array}{c|ccc} x_1y_1 & x_2y_2 & \ldots & x_1y_n \\ \hline \ldots & \ldots & \ldots & \ldots \\ x_my_1 & x_my_2 & \ldots & x_my_n \end{array}\right]$$

$$A = xy' = \left[\begin{array}{cc} x_1y_1 & x_1 (y_2, \ldots y_n) \\ y_1 \begin{pmatrix} x_2 \\ \vdots \\ x_m \end{pmatrix} & \begin{pmatrix} x_2y_2 \ldots\ldots x_2y_n \\ \text{----------} \\ \text{----------} \\ x_my_2 \ldots\ldots x_my_n \end{pmatrix} \end{array}\right]$$

This is in the form (2.7) where, $b = x_1 (y_2 \ldots y_n)$ is a $1 \times (n - 1)$ row vector, $c = y_1 (x_2 \ldots x_m)'$ is an $(m - 1) \times 1$ column vector, and

$$d = ba^{-1}c = y_1 \begin{pmatrix} x_2 \\ \ldots \\ x_m \end{pmatrix} \frac{1}{x_1y_1} x_1 (y_2 \ldots y_n) = \begin{pmatrix} x_2 \\ \ldots \\ x_m \end{pmatrix} (y_2 \ \ldots \ y_n)$$

is the outer product of an $(m - 1)$ – dimensional vector and an $(n - 1)$ – dimensional vector.

Remark : Note that (2.2) can also be written

$$A = \begin{pmatrix} I_r \\ ca^{-1} \end{pmatrix} (a \quad b) = \begin{pmatrix} a \\ c \end{pmatrix} a^{-1} (a \ b) = \begin{pmatrix} a \\ c \end{pmatrix} (I_r \ a^{-1}b)$$

This can be viewed as a generalization of the representation $A = uv'$ for an outer product of two vectors u, v.

Proof: If the first r rows of A are linearly independent and rank (A) = rank (a) = r in (2.6), then the last m − r rows of A are linear combinations of the first r rows. This means that we can write the last m − r rows of A as;

$$(c \quad d)_i = \sum_{j=1}^{r} T_{ij} (a \quad b)_j \quad \text{for } 1 \le i \le m - r$$

where, T_{ij} $(1 \le i \le m - r, 1 \le j \le r)$ are numbers. In terms of matrices,

$$(c \quad d) = T(a \quad b) = (Ta \quad Tb) \qquad \ldots (2.8)$$

where, T is $(m - r) \times r$.

The relation (2.8) implies c = Ta and hence $T = ca^{-1}$. This implies $Tb = ca^{-1} = d$ in (2.8), which completes the proof of Theorem 2.2.

Theorem 3 : Let $$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & ca^{-1}b \end{pmatrix} \qquad \ldots(2.9)$$

be an $m \times n$ matrix with r = rank (A) where a is $r \times r$ and invertible, as in Theorem 2.

Let $$G = \begin{pmatrix} a^{-1} & 0 \\ 0 & 0 \end{pmatrix} \qquad \ldots (2.10)$$

where, the "0" in (2.10) represent matrices of zeroes of dimension sufficient to make G an $n \times m$ matrix. Then G is a generalized inverse of A.

Proof : By equation (2.9) and (2.10).

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ ca^{-1} & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & ca^{-1}b \end{pmatrix}$$

where, I_r is the $r \times r$ unit matrix. This implies AGA = A since $d = ca^{-1} b$ by (2.9), so that G is a generalized inverse of A.

The two projections in this case are

$$AG = \begin{pmatrix} I_r & 0 \\ ca^{-1} & 0 \end{pmatrix} \quad \text{and} \quad GA = \begin{pmatrix} I_r & a^{-1}b \\ 0 & 0 \end{pmatrix}$$

Theorem 1 then says that $Ax = y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ can be solved for $y_1 \in R^r$, $y_2 \in R^{m-r}$ if and only if

$$AGy = \begin{pmatrix} I_r & 0 \\ ca^{-1} & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ ca^{-1}y_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

That is, if and only if $y_2 = ca^{-1}y_1$. In that case, the general solution of Ax = y for $x \in R^n$ is

$$\begin{aligned} x &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = Gy + (I_m - GA)\, z \\ &= \begin{pmatrix} a^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 & -a^{-1}b \\ 0 & I_{m-r} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= \begin{pmatrix} a^{-1}y_1 \\ 0 \end{pmatrix} + \begin{pmatrix} -a^{-1}bz_2 \\ z_2 \end{pmatrix} \end{aligned}$$

for arbitrary $z_2 \in R^{m-r}$.

Remark : This shows that any m × n matrix A has at least one generalized inverse G of the form (2.10). Since often many different linearly-independent sets of r rows can be permuted to the upper r rows and many different linearly-independent sets of r columns can be permuted into the first r column positions, a matrix A with rank (A) = r < n can have many different generalized inverse of this form.

Generalized Inverse :

General algorithm to find G (G = Generalized inverse).

(i) Suppose matrix A is given. Find any non-singular submatrix of order r_A. Denote it by M.

(ii) Find inverse of M and take transpose $(M^{-1})'$.

(iii) In A, replace each element of M by the corresponding elements of $(M^{-1})'$ i.e. $a_{ij} = m_{st}$, the $(s, t)^{th}$ element of M, then replace a_{ij} by $W^{t, s}$. The $(t, s)^{th}$ element of (M^{-1}), equivalent to $(s, t)^{th}$ element of $(M^{-1})'$.

(iv) Replace all other elements of A by zero.

(v) Transpose is the final resulting matrix.

(vi) The result is G, a generalized inverse of A.

SOLVED EXAMPLE

Example 2.1 : The matrix $A = \begin{bmatrix} 4 & 1 & 2 & 0 \\ 1 & 1 & 5 & 15 \\ 3 & 1 & 3 & 5 \end{bmatrix}$.

Solution : (i) Take $M = \begin{bmatrix} 1 & 5 \\ 1 & 3 \end{bmatrix}$

$\therefore \quad M^{-1} = \begin{bmatrix} -1.5 & 2.5 \\ 0.5 & -0.5 \end{bmatrix}$

$(M^{-1})' = \begin{bmatrix} -1.5 & 0.5 \\ 2.5 & -0.5 \end{bmatrix}$

$\therefore \quad G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1.5 & 2.5 \\ 0 & 5 & -0.5 \\ 0 & 0 & 0 \end{bmatrix}$

(ii) $$N = \begin{bmatrix} 1 & 15 \\ 1 & 5 \end{bmatrix}$$ (Submatrix)

$$N^{-1} = \begin{bmatrix} -0.5 & 1.5 \\ 0.1 & -0.1 \end{bmatrix}$$

$$(N^{-1})' = \begin{bmatrix} -0.5 & 0.1 \\ 1.5 & -0.1 \end{bmatrix}$$

The generalized inverse of A is

$$\therefore \quad G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -0.5 & 1.5 \\ 0 & 0 & 0 \\ 0 & 0.1 & -0.1 \end{bmatrix}$$

(iii) $$Z = \begin{bmatrix} 4 & 0 \\ 3 & 5 \end{bmatrix}$$

$$\therefore \quad Z^{-1} = \begin{bmatrix} 0.25 & 0 \\ -1.5 & 0.2 \end{bmatrix}$$

$$(Z^{-1})' = \begin{bmatrix} 0.25 & -1.5 \\ 0 & 0.2 \end{bmatrix}$$

$\therefore$ The generalized inverse of A is

$$G = \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1.5 & 0 & 0.2 \end{bmatrix}$$

Note : The procedure to obtain generalized inverse (In (iii) case), to replacing elements of M in A by the elements of (M^{-1}) and then in (V) take transpose.

2.4 Applications of Generalize Inverse in Statistics

Let X be an $n \times r$ matrix with $r > n$ and rank (X) = r. Then $X^{-1}X$ is invertible. If "observed values" $Y \in R^n$ can be "exactly fit" by the parameters $\beta \in R^r$, then

$$Y = X\beta, \quad Y \in R^n, \quad \beta \in R^r \qquad \ldots (2.11)$$

The matrix X cannot be invertible, since $r < n$. However, suppose that we want a general procedure to choose an arbitrary β in terms of Y. In that case, we can consider a generalized inverse of X. Specificially, G will be a generalized inverse of X if G is $r \times n$ and

$$XGX = X$$

Since X'X is invertible, an obvious choice is

$$G = (X'X)^{-1}X' \qquad \ldots (2.12)$$

since then $XGX = X(X'X)^{-1}X'X = X$. The two projections XG and GX are

$$GX = (X'X)^{-1}X'X = I_r$$

and

$$XG = X(X'X)^{-1} X' = H \qquad \ldots (2.13)$$

Note that both projections are symmetric : That is, $I_r = (I_r)'$ and $H = H'$.

In addition

$$GXG = ((X'X^{-1}X') X ((X'X)^{-1} X') = (X'X)^{-1} X' = G$$

That is, G is the unique Penrose inverse of the $n \times r$ matrix X.

Theorem 1 now says that $Y = X\beta$ can be solved exactly if and only if $(XG)Y = HY = Y$; that is, if and only if Y is in the range of the $n \times n$ projection H. Moreover, if $HY = Y$, then every solution of $X\beta = Y$ is of the form

$$\beta = GY = (X' X)^{-1} X' Y + (I_r - GX)\, z, \; z \in R^r$$

$$= (X' X)^{-1} X' Y \qquad \ldots (2.14)$$

since $GX = I_r$ by (2.13). In other words, if $Y = X\beta$ for some vector β, then the only solution β of $X\beta = Y$ for a given Y is given by (2.14).

It follows directly from (2.13) that X must be one-one : That is, if $X\beta_1 = X\beta_2$, then $GX\beta_1 = GX\beta_2 \Rightarrow \beta_1 = \beta_2$.

Suppose that Y as $X\beta$ that are observed with errors. That is, as

$$Y = X\beta + e \qquad \ldots (2.15)$$

where, $e = \{e_i\}$ are independent errors. Then we can consider the value of β that minimizes the sum of errors

$$\min_{\beta} \sum_{i=1}^{n} (Y_i - (X\beta)_i)^2 = \sum_{i=1}^{n} (Y_i - (X\hat{B})_i)^2 \qquad \ldots (2.16)$$

If we consider

$$\frac{\partial}{\partial \beta_j} \sum_{i=1}^{n} (Y_i - (X\beta)_i)^2 = \frac{\partial}{\partial \beta_j} \sum_{i=1}^{n} \left(Y_i - \sum_{a=1}^{r} X_{ia}\beta_a \right)^2$$

$$= -2 \sum_{i=1}^{n} \left(Y_i - \sum_{a=1}^{r} X_{ia}\beta_a \right) X_{ij} = 0$$

This implies

$$\sum_{a=1}^{r} \sum_{i=1}^{n} X_{ij}\, X_{ia}\beta_a = \sum_{i=1}^{n} X_{ij} Y_j, \quad 1 \le j < r$$

which can be written in a more compact form as

$$(X' X)\, \beta = X' Y \qquad \ldots (2.17)$$

Since we are assuming that X' X is invertible, (2.17) implies

$$\hat{B} = (X' X)^{-1} X' Y = GY$$

for G in (2.12). That is, the *least-squares* solution of (2.16) for β is given by;

$$\beta = \hat{\beta} = GY$$

where, G is the Penrose inverse of the $n \times r$ matrix X.

2.5 Moore-Penrose Generalized Inverse

Moore (1920) and Penrose (1955) reduced the infinitely many generalized inverse to one unique solution by imposing four reasonable conditions, all met by the standard inverse. If

1. general condition : AGA = A,
2. reflective condition : GAG = G,
3. normalized condition : (AG)' = GA, and
4. reverse normalized condition : (GA)' = AG

then this G matrix is unique matrix.

Note : A moore penrose generalized inverse is denoted by A^+ i.e. $G = A^+$.

A matrix that satisfies the first two conditions is called a "reflexive" or "weak" generalized inverse and is order dependent. A matrix that satisfies the first three conditions is called a "normalized" generalized inverse. A matrix that satisfies the first and fourth conditions is called a "minimum norm" generalized inverse.

The Moore-Penrose generalized inverse is also easy to calculate using QR factorization. QR factorization takes the input matrix A and factors it into the product of an orthogonal matrix, Q and a matrix, R, which has a triangular leading square matrix (r) followed by rows of zeros corresponding to the difference in the rank and dimension in A :

$$A = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

This factorization is implemented in every professional level statistical package. The Moore-Penrose generalized inverse is produced by :

$$C = [r^{-1}\ 0]\ Q'$$

where, 0 is the transpose of the zero matrix of the R matrix required for conformability.

[**Note :** AG is always idempotent (GAGA = G(AGA) = GA) and rank (AG) = rank (A). These results hold whether A is singular or not.]

Note : To find a unique generalized inverse (Moore Penrose generalized inverse) we use full rank factorization method (discuss in the later section). In this method A matrix $A_{p \times q}$ factorize in two full rank (row rank) and (column rank) matrices i.e. A = KL where K and L have full column and row rank respectively equals to rank of matrix A (r) then M (Moore Penrose generalized inverse) is

$$M = L' (K'AL')^{-1} K'$$

here K'K and LL' are non-singular matrices.

Example : Let $A = \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & 2 & 2 \\ -1 & 4 & 5 & 3 \end{bmatrix}$

then $A = KL = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$

where matrix L is obtained by taking non-zero rows of row echelon form of matrix A and matrix K is obtained by taking non-zero Pivotal columns of matrix A.

$$\therefore \quad K'AL' = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & 2 & 2 \\ -1 & 4 & 5 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & -12 \\ -12 & 60 \end{bmatrix}$$

$$\therefore \quad M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \left(\frac{1}{36}\right) \begin{bmatrix} 10 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 4 \end{bmatrix}$$

$$= \frac{1}{18} \begin{bmatrix} 5 & 2 & -1 \\ 1 & 1 & 1 \\ -4 & -1 & 2 \\ 6 & 3 & 0 \end{bmatrix}$$

Here M is known as Moore-Penrose inverse of A. (A^+). If $A_{p \times q}$ then $M_{q \times p}$ matrix. When A is non-signular then M is regular inverse i.e. $M = A^{-1}$.

Theorem 5 : For any matrix-A : $m \times n$ and any g-inverse we have

(i) A^-A and AA^- are idempotent.

(ii) rank (A) = rank (AA^-) = rank (A^-A).

(iii) rank (A) ≤ rank (A^-).

Proof :

(a) Using the definition of g-inverse.

$$(A^-A)(A^-A) = A^-(AA^-A) = A^-A$$

$\Rightarrow$ A^-A is idempotent, similarly we can prove that AA^- is also idempotent.

(b) We know that rank $(AB) \leq$ rank (A) or rank (B).

$$\text{rank } (A) = \text{rank } (AA^-A) \leq \text{rank } (A^-A) \leq \text{rank } (A),$$

that is, rank (A^-A) = rank (A). Analogously, we see that rank (A) = rank (AA^-).

(c) rank (A) = rank $(AA^-A) \leq$ rank $(AA^-) \leq$ rank (A^-).

Remarks : Let A be an $m \times n$-matrix. Then

(i) A is regular matrix $\Rightarrow A^+ = A^{-1}$.

(ii) $(A^+)^+ = A$.

(iii) $(A^+)' = (A')^+$.

(iv) rank (A) = rank (A^+) = rank (A^+A) = rank (AA^+).

(v) A an orthogonal matrix $\Rightarrow A^+ = A$.

(vi) If rank $(A_{m \times n}) = m \Rightarrow A^+ = A'(AA')^{-1}$ and $AA^+ = I_m$.

(vii) If rank $(A_m \times n) = n$ rank $(A) : m \times n = n \Rightarrow A^+ = (A'A)^{-1}A'$ and $A^+A = I_n$.

(viii) If $P : m \times m$ and $Q : n \times n$ are orthogonal $\Rightarrow (PAQ)^+ = Q^{-1} A^+ P^{-1}$.

(ix) $(A'A)^+ = A^+(A')^+$ and $(AA')^+ = (A')^+ A^+$.

(x) $A^+ = (A'A)^+ A' = A' (AA')^+$.

2.6 Existence of g-inverse by Different Methods

First observe that if A is square and non-singular, then A^{-1} is the only g-inverse of A. Also, if the matrix B has full column rank (i.e., the column of B are linearly independent), then the left inverse of B are its g-inverses. A similar remark applies to matrices of full row rank. We now turn to the existence of g-inverse of an arbitrary matrix.

1. Using rank factorization : Any matrix having a rank factorization. Thus if A is an $m \times n$ matrix of rank r, then there exist matrices B and C of rank r and of order $m \times r$ and $r \times n$ respectively such that $A = BC$. Clearly, B admits a left inverse, say B^- and C admits a right inverse, say C^-. Then $G = C^- B^-$ is a g-inverse of A since $AGA = BCC^- B^- BC = BC = A$. Note that if we choose B^- to be a least-squares g-inverse (in fact such a g-inverse must be B^+ since B has full column rank) then $G = C^- B^-$ is a least-squares g-inverse of A. Similarly choosing C^- to be a minimum-norm g-inverse (which must be C^+) we obtain a minimum-norm g-inverse of A. Combining these observations, $A^+ = C^+ B^+$ is the Moore-Penrose inverse of A (unique g-inverse).

2. Using the rank canonical form : If A is $m \times n$ of rank r, then there exist non-singular matrices P and Q of order $m \times m$ and $n \times n$ respectively, such that

$$A = P \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q.$$

It can be verified that for any U, V, W of appropriate dimensions,

$$\begin{bmatrix} I_r & U \\ V & W \end{bmatrix}$$

is a g-inverse of $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

Then $$G = Q^{-1} \begin{bmatrix} I_r & U \\ V & W \end{bmatrix} P^{-1}$$

is a g-inverse of A.

This approach makes it evident that if A is not a square non-singular matrix, then it admits infinitely many g-inverses, since each choice of U, V, W leads to a different g-inverse. We can also deduce the following fact : if A has rank r and if $r < k < \min \{m, n\}$ then A admits a g-inverse of rank k. To see this, choose U and V to be zero and W to be a matrix of rank k. To see this, choose U and V to be zero and W to be a matrix of rank $k - r$. In particular, any square matrix admits a non-singular g-inverse.

3. A computational approach : Let A be an $m \times n$ matrix of rank r. Then A has an $r \times r$ non-singular submatrix. We assume that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where, A_{11} is $r \times r$ and non-singular. Since the rank of A is r, the first r columns of A form a basis for the column space and hence $A_{12} = A_{11}X$ and $A_{22} = A_{21}X$ for some matrix X. Let,

$$G = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

where, the zero blocks are chosen so as to make G and $n \times m$ matrix. Then a simple computation shows that

$$\begin{aligned} AGA &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{11}X \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}X \end{bmatrix} \\ &= A \end{aligned}$$

As an example, consider the matrix

$$A = \begin{bmatrix} 1 & 7 & 1 & 4 \\ 1 & 2 & 0 & 1 \\ 0 & 5 & 1 & 3 \end{bmatrix}$$

of rank 2. For illustration, we pick the non-singular 2×2 submatrix of A formed by rows 2, 3 and columns 2, 4. Replacing the matrix by its inverse transpose and the other entries by zero, gives the matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -5 \\ 0 & -1 & 0 & 2 \end{bmatrix}$$

Transposing the above matrix results in a g-inverse of A,

$$G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \\ 0 & -5 & 2 \end{bmatrix} = A^{-}$$

4. By induction method : Let A be an $m \times n$ matrix of rank r. We assume, without loss of generality, that $r < n \leq m$. Then A must have a column, again take it to be the last column for convenience, which is a linear combination of the rest. Write A as an augmented matrix $A = [B, x]$ where, x is the last column of A. Proceeding by induction on $m + n$, we may assume that B has a g-inverse say B^{-}. Suppose

$$G = \begin{bmatrix} B^{-} \\ 0 \end{bmatrix}.$$

Observe that claim G is a g-inverse of A.

$$AGA = BB^{-} A = BB^{-} [B, x] = [B, BB^{-} x].$$

Since, x is in the column space of B, $x = By$ for some y. Then $BB^{-} x = BB^{-} By = By = x$. It follows that $AGA = A$ and the claim is proved.

More generally, suppose A is partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where, A_{11} is of rank r. Then

$$G = \begin{bmatrix} A_{11}^{-} & 0 \\ 0 & 0 \end{bmatrix}$$

is a g-inverse of A for any choice of a g-inverse A_{11}^{-} of A_{11}. (Here the zero blocks in G are having order to make it an $n \times m$ matrix.)

5. An operator theoretic approach : Consider the map f_A defined from $C^{m \times n}$ to itself given by $f_A(B) = AB^*A$. We show that A is in the range of f_A. On $C^{m \times n}$ we have the usual inner product $< X, Y > = \text{tr}(XY^*)$. Suppose Z is orthogonal to the range of f_A. Then $\text{tr}(AB^* AZ^*) = 0$ for all $B \in C^{m \times n}$ and hence $\text{tr}(B^* AZ^* A) = 0$ for all $B \in C^{m \times n}$. It follows that $AZ^* A = 0$. Thus $AZ^* AZ^* = 0$ and hence $(AZ^*)^2 = 0$. Thus all eigenvalues of AZ^* are zero and therefore $\text{tr}(AZ^*) = 0$. Thus Z is orthogonal to A. It follows that any matrix which is orthogonal to the range of f_A is orthogonal to A and hence A is in the range of f_A. Thus there exists $H \in C^{m \times n}$ such that $f_A(H) = AH^*A = A$. Then $G = H^*$ is a g-inverse of A.

6. A functional approach : Subspaces S and T of a vector space V are said to be complementary if $V = S + T$ and $S \cap T = \{0\}$. If A is an $m \times n$ matrix then we denote by $N(A) \subset C^n$ the null space of A and by $R(A) \subset C^m$ the range space of A. Let S be a subspace of C^m complementary to N(A) and let T be a subspace of C^m complementary to R(A). The matrix A defines a linear map from C^n to C^m. Let A_1 be the restriction of the map to S. Then it is easily seen that A_1 is a one-to-one map from S onto R(A). Thus $A_1^{-1} x$, $x \in R(A)$ is well-defined. Define $G : C^m \to C^n$ as $Gx = A_1^{-1} x$ if $x \in R(A)$, $Gx = 0$ if $x \in T$ and extend linearity to C^m. ($x \in C^m$ can be uniquely written as $x = y + z$, where $y \in R(A)$, $z \in T$. Then set $Gx = Gy + Gz$.) We denote the matrix of G with respect to the standard basis by G itself. Then it is easy to check that $AGA = A$ and hence we have got a g-inverse of A.

We remark that if $S = R(A^*)$ and $T = S^{\perp} = N(A^*)$ then the g-inverse constructed above is the Moore-Penrose inverse.

7. Using the spectral theorem : Suppose A is an $n \times n$ hermitian matrix. By the spectral theorem, there exists a unitary matrix U such that

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} U^*$$

where, D is the diagonal matrix with the non-zero eigenvalues of A along its diagonal. Then

$$G = U^* \begin{bmatrix} U^{-1} & X \\ Y & Z \end{bmatrix} U$$

is a g-inverse of A for arbitrary X, Y, Z. Thus any hermitian matrix admits a g-inverse. Now for an arbitrary $m \times n$ matrix A, it can be verified that $G = (A^*A)^- A^*$ is a g-inverse of A for an arbitrary g-inverse $(A^*A)^-$ of the hermitian matrix A^*A.

8. Using the singular value decomposition : If A is an $m \times n$ matrix then there exist unitary matrices U and V such that

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V$$

where D is the diagonal matrix with the singular values of A along its diagonal. This is the well-known singular value decomposition of A. Then

$$G = V^* \begin{bmatrix} D^{-1} & X \\ Y & Z \end{bmatrix} U^*$$

is a g-inverse of A for arbitrary X, Y, Z. If X = 0 then G is a least-square g-inverse, if Y = 0 then G is a minimum-norm g-inverse and finally, if X, Y and Z are all null matrices then G is the Moore-Penrose inverse.

9. A proof using linear equations : If A is an m × n matrix then recall that the column space of A* equals the column space of A*A. Therefore, the equation A*AX = A* is consistent. Thus there exists an n × m matrix G such that A*AG = A*. Now

$$AGA = G^* A^* AGA = G^* A^* A = A$$

and hence G is a g-inverse of A.

Theorem 6 : For any matrix A we have

$$A'A = 0 \text{ if and only if } A = 0.$$

Proof :

(a) $A = 0 \Rightarrow A'A = 0$.

(b) Let $A'A = 0$ and let $A = (a_{(1)}, \ldots, a_{(n)})$ be the columwise presentation. Then

$$A'A = \left(a'_{(i)}\, a_{(j)}\right) = 0,$$

so that all the elements on the diagonal are zero : $a'_{(i)}\, a_{(i)} = 0 \Rightarrow a_{(i)} = 0$ and $A = 0$.

Theorem 7 : Let $X \neq 0$ be an m × n-matrix and A an n × n-matrix. Then

$$X'XAX'X = X'X \Rightarrow XAX'X = X \text{ and } X'XAX' = X'$$

Proof : As $X \neq 0$ and $X'X \neq 0$, we have,

$$\begin{aligned} X'XAX'X - X'X &= (X'XA - I)\, X'X = 0 \\ &\Rightarrow (X'XA - I) = 0 \\ \Rightarrow 0 &= (X'XA - I)\,(X'XAX'X - X'X) \\ &= (X'XAX' - X')\,(XAX'X - X) = Y'Y \end{aligned}$$

so that if Y = 0 and hence XAX'X = X.

Corollary : Let $X \neq 0$ be an m × n-matrix and A be n × n-matrices.

Then $AX'X = BX'X \Leftrightarrow AX' = BX'$

2.7 Singular Value Decomposition

Introduction :

Any real $m \times n$ matrix can be factored as

$$A = UDV^T$$

where, U is an $m \times m$ orthogonal matrix whose columns are the eigenvectors of AA^T, V is an $n \times n$ orthogonal matrix whose columns are the eigenevectors of A^TA and D is an $m \times n$ diagonal matrix of the form

$$D = \begin{bmatrix} d_1 & \dots & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ 0 & 0 & \dots & \dots \\ : & : & : & : \\ 0 & 0 & d_r & 0 \\ : & : & : & : \\ 0 & 0 & 0 & \dots 0 \end{bmatrix}$$

with $d_1 \geq d_2 \geq \dots \geq d_r > 0$ and r = rank (A). In the above, $d_1, \dots, d_r$ are the square roots of the eigenvalues of A^TA. They are called the *singular values* of A.

First "solve" the system Ax = b for all matrices A and vectors b. Also solve the system using a numerical algorithm and solve the system in a reasonably efficient manner. For instance, we do not want to compute A^{-1} using determinants.

There are three situations arise.

(a) If A is square and invertible, we want to have the solution $x = A^{-1}b$.

(b) If A is under constrained, we want the entire set of solutions.

(c) If A is overconstrained, we could find least-squares solution. In other words we want that x which minimizes the error $\|Ax - b\|$. Geometrically, $A\bar{x}$ is the point in the column space of A closest to B.

Gaussian elimination is reasonably efficient, but it is not numerically very good. In particular, elimination does not deal with singular matrices. The method is not designed for overconstrained systems. Even for underconstrained systems, the method requires extra calculation.

The poor numerical character of elimination can be seen in a two different ways. First, the elimination process assumes a non-singular matrix. Two sets below, are linearly independent vectors.

$$\left\{\begin{pmatrix}1\\0\\0\end{pmatrix} \cdot \begin{pmatrix}0\\1\\0\end{pmatrix} \cdot \begin{pmatrix}0\\0\\1\end{pmatrix}\right\} = \left\{\begin{pmatrix}1.01\\1.00\\1.00\end{pmatrix} \cdot \begin{pmatrix}1.00\\1.01\\1.00\end{pmatrix} \cdot \begin{pmatrix}1.00\\1.00\\1.01\end{pmatrix}\right\}$$

The first set is independent, while the second set is "almost dependent" set of vectors.

Second, elimination-based methods work like LU decomposition, which represents the coefficient matrix A as a matrix product LDU, where L and U are respectively lower and upper diagonal and D is diagonal. One solves the system Ax = b by solving (via backsubstitution) Ly = b and Ux = D^{-1}y. If A is singular, then D will contain near-zero entries on its diagonal.

Singular values decomposition is a powerful technique for dealing with sets of equations or matrices that are either singular or very close to singular. In many cases where Gaussian elimination and LU decomposition fail to give satisfactory results, SVD will not only diagnoise the problem but also give you a useful numerical answer. It is also the method of choice for solving most linear least-squares problems.

2.8 SVD (Singular Value Decomposition) Algorithm

It is useful to establish a relation with Gaussian elimination, which reduces a matrix A by a series of row operations that gives non-zero in entries in columns of A. Row operations imply pre-multiplying the matrix A. They are all collected together in the matrix L^{-1} where A = LDU.

But in SVD, non-zero entries are occur in both rows and columns of A. Thus whereas Gaussian elimination only reduces A using pre-multiplication. SVD uses both pre and post-multiplication. As a result, SVD can at each stage rely on orthogonal matrices to perform its reductions on A. By using orthogonal matrices, SVD reduces the risk of magnifying noise and errors. The pre-multiplication matrices are gathered together in the matrix U^T, while the post-multiplication matrices are gathered together in the matrix V.

There are two phases of SVD decomposition algorithm :

(i) SVD reduces A to bidiagonal form using a series of orthogonal transformations. This phase is deterministic and has a running time that depends only on the size of the matrix A.

(ii) SVD removes the superdiagonal elements from the bidigonal matrix using orthogonal transformation. This phase is iterative but converges quickly.

Let us take a slightly closer look at the first phase. Set 1 in this phase creates two orthogonal matrices U_1 and V_1 such that

$$U_1 A = \begin{Bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{Bmatrix}$$

$$U_1 A V_1 = \begin{Bmatrix} a''_{11} & a''_{12} & 0 & \cdots & 0 \\ 0 & & & & \\ \vdots & & B'' & & \\ 0 & & & & \end{Bmatrix}$$

If A is m × n then B" is (m − 1) × (n − 1). The next step of this phase recursively works on B" and so forth, until orthogonal matrices U_1, …, U_{n-1}, V_1, …, V_{n-2} are produced such that $U_{n-1} \ldots U_1 A V_1 \ldots V_{n-2}$ is bidiagonal (assume m ≥ n).

In both phases, the orthogonal transformation are constructed from Householder matrices.

SOLVED EXAMPLE

Example 2.2 : Find the SVD of A, UDV^T, where $A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$.

First we compute the singular values d, by finding the eigenvalues of AA^T.

$$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$$

The characteristics polynomial is $\det(AA^T - \lambda I) = \lambda^2 - 34\lambda + 225 = (\lambda - 25)(\lambda - 9)$, so the singular values are $d_1 = \sqrt{25} = 5$ and $d_2 = \sqrt{9} - 3$.

Now we find the right singular vectors (the columns of V) by finding an orthonormal set of eigenvectors of A^TA. It is also possible to proceed by finding the left singular vectors (columns of U) instead. The eigenvalues of A^TA are 25, 9 and 0 and since A^TA is symmetric. We know that the eigenvectors will be orthogonal.,

For $\lambda = 25$, we have

$$A^TA - 25I = \begin{pmatrix} -12 & +12 & +2 \\ +12 & -12 & -2 \\ +2 & -2 & -17 \end{pmatrix}$$

which row-reduces to $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. A unit-length vector in the kernel of that matrix is $v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$.

For $\lambda = 9$ we have $A^TA - 9I = \begin{pmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & 2 & 1 \end{pmatrix}$ which row-reduces to $\begin{pmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{pmatrix}$.

A unit-length vector in the kernel is $v_2 = \begin{pmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{pmatrix}$.

For the last eigenvector, we could compute the kernel of A^TA or find a unit vector perpendicular to v_1 and v_2. To be perpendicular to $v_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ we need a = b. Then the condition that $v_2^T v_3 = 0$ becomes $2a/\sqrt{18} + 4c/\sqrt{18} = 0$ or a = 2c. So $v_3 = \begin{pmatrix} a \\ a \\ a/2 \end{pmatrix}$ and for it to be unit-length we need a = 2/3 so $v_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix}$.

So at this point we know that

$$A = UDV^T = U \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & -2/3 & -1/3 \end{pmatrix}$$

Finally, we can compute U by the formula $du_i = Av_i$ or $u_i = \frac{1}{d} Av_i$. This gives $A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$.

$\therefore$ The SVD of a matrix A is

$$A = UDV^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & -2/3 & -1/3 \end{pmatrix}$$

2.9 Full Rank Factorization

Introduction : Triangular factorization of matrices play an important role in solving linear systems. It is known that the LDU factorization is unique for square nonsingular matrices and for full row rank rectangular matrices. In any other case, the LDU factorization is not unique and the orders of L, D and U are greater than the rank of the orginal matrix.

Consider matrices $A \in \mathbb{R}^{n \times m}$ with rank (A) = r $\leq$ min {n, m}, where the LDU factorization of A is not unique. For this kind of matrices, it is useful to consider the *full rank factorization of A,* that is, a decomposition in the form A = FG with $F \in \mathbb{R}^{n \times r}$ $\Rightarrow$ $G \in R^{r \times m}$ and rank (F) = rank (G) = r. The full rank factorization of any non-zero matrix is not unique. In addition, if A = FG is a full rank factorization of A, then any other full rank factorization can be written in the form $A = (FM^{-1})(MG)$, where $M \in R^{r \times r}$ is a non-singular matrix.

If the full rank factorization of A is given by A = LDU, where $L \in R^{n \times r}$ is in lower echelon form, D = diag ($d_1, d_2, \ldots, d_r$) is non-singular and $U \in \mathbb{R}^{r \times m}$ is in upper echelon form, then this factorization is called a *full rank factorization in echelon form of A.*

In this paper we give a method to obtain a full rank factorization of a rectangular matrix and we study when this decomposition can be in echelon form. Moreover, if the factorization in echelon form exists, we can prove that it is unique. Finally, applying the full rank factorization in echelon form, we give a simple proof of the Flanders theorem [4] for matrices $A \in \mathbb{R}^{n \times r}$ and $B \in R^{r \times n}$ with rank (A) = rank (B) = r, as well as of the converse result.

The full rank factorization of matrices has different applications, for example, in control theory to obtain minimal realizations of polynomial transfer matrices by using the Silverman-Ho algorithm. In numerical analysis to obtain a Cholesky full rank factorization or to extend the QR factorization to rectangular matrices without full rank. In matrix analysis to obtain the non-zero eigenvalues and their associated eigenvectors or the singular values of a matrix.

Moreover, this factorization allows us to characterize some particular classes of matrices because not all matrices have a full rank factorization in echelon form.

SOLVED EXAMPLE

Example 2.3 : Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ -1 & -1 & 2 & -2 & -5 \\ 2 & 3 & -1 & 0 & -2 \\ 2 & 3 & 5 & -1 & 2 \\ 0 & 1 & 4 & 0 & -1 \end{bmatrix}$$

Since A has no zero rows, we have $\bar{A} = A$. Then by applying the first iteration of the Gauss elimination process. ($\bar{A}$ denote the matrix after transformation).

$$A_{(1)} = E^{(1)}_{4,1}(-2)\, E^{(1)}_{3,1}(-2)\, E^{(1)}_{2,1}(A)\, \bar{A} = \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 1 & 9 & -5 & -8 \\ 0 & 1 & 4 & 0 & -1 \end{bmatrix}$$

By deleting the zero row we obtain

$$A_{(1)} = F^{(2)}_5 \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 1 & 9 & -5 & -8 \\ 0 & 1 & 4 & 0 & -1 \end{bmatrix} = F^{(2)}_5 \bar{A}_{(1)}.$$

Now the second iteration of the Gauss process is applied to $\bar{A}_{(1)}$ giving

$$A_{(2)} = E^{(2)}_{4,2}(-1)\, E^{(2)}_{3,2}(-1)\, \bar{A}_{(1)} = \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 0 & 6 & -1 & 4 \\ 0 & 0 & 1 & 4 & 11 \end{bmatrix}$$

This matrix has no zero row, so $\bar{A}_{(2)} = A_{(2)}$ and following with the third iteration of the Gauss elimination process we obtain

$$A_{(3)} = E^{(3)}_{4,3}(-1/6)\, \bar{A}_{(2)} = \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 0 & 6 & -1 & 4 \\ 0 & 0 & 0 & 25/6 & 31/3 \end{bmatrix}$$

which can be written as

$$A_{(3)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 25/6 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 0 & 1 & -1/6 & 2/3 \\ 0 & 0 & 0 & 1 & 62/25 \end{bmatrix} = DU$$

where $D \in \mathbb{R}^{4\times 4}$ is a nonsingular diagonal matrix and $U \in \mathbb{R}^{4\times 5}$ is an upper echelon matrix with rank (U) = 4. Finally, we have that the full rank factorization in echelon form of A is

$$A = \left(E^{(1)}_{2,1}(-1)\, E^{(1)}_{3,1}(2)\, E^{(1)}_{4,1}(2)\, F^{(2)}_{5}(1)\, E^{(2)}_{3,2}(1)\, E^{(2)}_{4,2}(1)\, E^{(3)}_{4,3}(1/6)\right) DU = LDU,$$

where $L \in \mathbb{R}^{5\times 4}$ is the following lower echelon matrix with rank (L) = 4,

$$L = E^{(1)}_{2,1}(-1)\, E^{(1)}_{3,1}(2)\, E^{(1)}_{4,1}(2)\, F^{(2)}_{5}(1)\, E^{(2)}_{3,2}(1)\, E^{(2)}_{4,2}(1)\, E^{(3)}_{4,3}(1/6) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 1/5 & 1 \end{bmatrix}$$

Remark : Note that it is not possible to apply the qusi elimination process to A without pivoting. Therefore, the full rank factorization of A obtained applying this method is not a full rank factorization in echelon form. Specially, if we apply the quasi-Neville elimination process to A we have that

$$A_{(1)} = E_2^{(1)}(1)\,E_3^{(1)}(-2)\,E_4^{(1)}(-1)\,\bar{A} = \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 0 & 6 & -1 & 4 \\ 0 & 1 & 4 & 0 & -1 \end{bmatrix} = F_5^{\{2\}}\,\bar{A}_{(1)}.$$

where, $$\bar{A}_{(1)} = \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 0 & 6 & -1 & 4 \\ 0 & 1 & 4 & 0 & -1 \end{bmatrix}$$

Now, it is not possible to apply the quasi elimination process to $\bar{A}_{(1)}$ without interchanging the rows. Finally, the full rank factorization of A that we obtain applying this method is

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 6 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 & 2 & 5 \\ 0 & 1 & 3 & -4 & -12 \\ 0 & 0 & 1 & 4 & 11 \\ 0 & 0 & 0 & -25 & -62 \end{bmatrix} = FG$$

which is not in echelon form. Therefore, as we comment at the beginning of this section, we can conclude that the quasi-Gauss elimination process is more general than the quasi elimination process.

Full Rank Factorization in Echelon Form : In this section we derive a necessary and sufficient condition for a matrix to have full rank decomposition in echelon form. Moreover, we prove that if the full rank factorization in echelon form exists then it is unique.

Remark : The full rank factorization in echelon form allows us to know, from the row of L with a leading 1, the linear independent rows beginning from the top of A. Therefore, if we have two different full rank factorization in echelon form, then one row can be independent or dependent of the rows above it at the same time, which is absurd. Taking into account this comment, we can give the following results.

Theorem 8 : Let $A \in \mathbb{R}^{n \times m}$ be a matrix with rank (A) = $r \le \min\{n, m\}$. If A has a full rank factorization in echelon form, then this factorization is unique.

Proof : Suppose that there exist two full rank factorizations in echelon form of A

$$L = L_1D_1U_1 = L_2D_2U_2,$$

where $L_1, L_2 \in \mathbb{R}^{n \times r}$ are lower echelon matrices with rank (L_1) = rank (L_2) = r, D_1 = diag ($d_{11}, d_{12}, \ldots, d_{1r}$) and D_2 = diag ($d_{21}, d_{22}, \ldots, d_{2r}$) nonsingular matrices and $U_1, U_2 \in \mathbb{R}^{r \times m}$ are upper echelon matrices with rank (U_1) = rank (U_2) = r.

We have that necessarily $L_1 = L_2 = L$. Then,

$$A = LD_1U_1 = LD_2U_2$$

Since L can be written, in a unique way, as $L = FL_{11}$, where F is a reduced lower echelon matrix and L_{11} is a unit lower triangular matrix, then

$$A = FL_{11}D_1U_1 = FL_{11}D_2U_2$$

From this equality $L_{11}D_1U_1$ and $L_{11}D_2U_2$ are two different factorizations with no pivoting of the submatrix $A_1 \in \mathbb{R}^{r \times m}$ formed by the first r linearly independent rows of A, which it is not possible. Therefore, $D_1 = D_2$ and $U_1 = U_2$.

Remark : We have proven that the full rank factorization in echelon form of A exists if the upper echelon form of the first r linearly independent rows can be obtained with no pivoting, and in this case we have obtained the factorization. We want to point that if $\bar{A}_1$ is the submatrix formed by any r linear independent rows of A and the full rank factorization in echelon form of A exists, then it can be obtained from $\bar{A}_1$ if the following conditions hold :

(i) The matrix F_1, such that $A = F_1\bar{A}_1$, is in lower echelon form.

(ii) The echelon form of $\bar{A}_1$ can be obtained without interchange of rows.

Find Mp-g inverse of following matrices :

SOLVED EXAMPLES

Example 2.4 : $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}_{2\times 2}$

Solution : The given matrix is $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}_{2\times 2}$. To find Mp-g inverse we produce as follows :

$R_2 \to R_1$ $\quad A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}_{2\times 2}$

$L = [1 \quad 2]_{1\times 2}$

Now let $\quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$R_2 + R_1 \Rightarrow \quad I = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

$K = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{2\times 1}$ where we have

$M = L' (K'AL')^{-1} K^1$

$$(K'AL') = [1 \quad 1] \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 10$$

$$(K'AL')^{-1} = \left[\frac{1}{10}\right]$$

$$L'(K'AL')^{-1} K' = \left[\frac{1}{2}\right]\left[\frac{1}{10}\right] [1 \quad 1]$$

$$M = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} \\ \frac{2}{10} & \frac{2}{10} \end{bmatrix}$$

$\therefore$ Mp inverse of matrix A is

$$A = \begin{bmatrix} 1/10 & 1/10 \\ 2/10 & 2/10 \end{bmatrix} \qquad \because AMA = A$$

Example 2.5 : $B = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2 & 2 & 4 \\ 3 & 6 & 1 & 2 \\ 4 & 8 & 2 & 4 \end{bmatrix}$

Solution : The given matrix is $B = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2 & 2 & 4 \\ 3 & 6 & 1 & 2 \\ 4 & 8 & 2 & 4 \end{bmatrix}$

$R_2 - R_1$
$R_3 - 3R_1$
$R_4 - 4R_1$

$$B = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & -8 & -16 \\ 0 & 0 & -10 & -20 \end{bmatrix}$$

$R_3 - 8R_2$
$R_4 - 10R_2$

$$\therefore B = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$(-R_2)$

$$B = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Now, $$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - R_2 \quad I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$R_3 + 8R_1$
$R_4 + 10R_2$ $$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & -8 & 0 & 0 \\ 0 & -10 & 0 & 0 \end{bmatrix}$$

$R_2 + R_1$
$R_3 + 3R_1$
$R_4 + 4R_1$ $$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & -8 & 0 & 0 \\ 4 & -10 & 0 & 0 \end{bmatrix}$$

$\therefore$ $$K = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 3 & -8 \\ 4 & -10 \end{bmatrix} \qquad A = K \cdot L$$

$$(K'AL') = \begin{bmatrix} 1 & 1 & 3 & 4 \\ 0 & -1 & -8 & -10 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2 & 2 & 4 \\ 3 & 6 & 1 & 2 \\ 4 & 8 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \\ 6 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 27 & 54 & 16 & 30 \\ -65 & -130 & -30 & -62 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \\ 6 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 375 & 80 \\ -775 & -150 \end{bmatrix}$$

$$\det (K'AL') = 5750$$

$$(K'AL') = \frac{1}{5750} \begin{bmatrix} -150 & -80 \\ 775 & 375 \end{bmatrix}$$

We have, $M = L'(K'AL')^{-1} K'$

$$= \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \\ 6 & 2 \end{bmatrix} \frac{1}{5750} \begin{bmatrix} -150 & -80 \\ 775 & 375 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 & 4 \\ 0 & -1 & -8 & -10 \end{bmatrix}$$

$$= \frac{1}{5750} \begin{bmatrix} -150 & -80 \\ -300 & -160 \\ 325 & 135 \\ 650 & 270 \end{bmatrix} \begin{bmatrix} 1 & 1 & 3 & 4 \\ 0 & -1 & -8 & -10 \end{bmatrix}$$

$$M = \frac{1}{5750} \begin{bmatrix} -150 & -70 & 190 & 200 \\ -300 & -140 & 380 & -400 \\ 325 & 190 & -105 & -50 \\ 650 & 380 & -210 & -100 \end{bmatrix}$$

$\therefore$ M is Mp-g inverse of A.

Example 2.6 : $A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & 1 & 7 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & 1 & 7 \end{bmatrix}$

$R_2 + 3R_1 \rightarrow A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 7 & 16 \end{bmatrix}$

$R_2\left(\frac{1}{7}\right) \rightarrow A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 16/17 \end{bmatrix}$

$L = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 16/17 \end{bmatrix}$

Now, $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$7R_2 \rightarrow I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$\xrightarrow{R_2 - 3R_1} I = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$K = \begin{bmatrix} 1 & 0 \\ -3 & 7 \end{bmatrix}$$

$$A = L \cdot K$$

$$(K'AL') = \begin{bmatrix} 1 & -3 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -3 & 1 & 7 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 16/17 \end{bmatrix} = \begin{bmatrix} -46 & -295/7 \\ 140 & 119 \end{bmatrix}$$

$$|K'AL'| = 426$$

$$(K'AL')^{-1} = \frac{1}{426} \begin{bmatrix} -45 & 295/7 \\ 140 & 119 \end{bmatrix}$$

Now $$M = L' (K'AL')^{-1} K'$$

$$= \frac{1}{426} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 16/7 \end{bmatrix} \begin{bmatrix} 119 & 295/7 \\ -140 & -46 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 7 \end{bmatrix}$$

$$= \frac{1}{426} \begin{bmatrix} 119 & 295/7 \\ 98 & 268/7 \\ 37 & 149 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 7 \end{bmatrix} = \frac{1}{426} \begin{bmatrix} 119 & -62 \\ 98 & -26 \\ 37 & 932 \end{bmatrix}$$

Important Points

- If A is an $m \times n$ matrix, then G is a *generalized inverse* of A is an $n \times m$ matrix G with

$$AGA = A$$

- General algorithm to find G (G = Generalized inverse).
  (i) Suppose matrix A is given. Find any non-singular submatrix of order r_A. Denote it by M.
  (ii) Find inverse of M and take transpose $(M^{-1})'$.
  (iii) In A, replace each element of M by the corresponding elements of $(M^{-1})'$ i.e. $a_{ij} = m_{st}$, the $(s, t)^{th}$ element of M, then replace a_{ij} by $W^{t, s}$. The $(t, s)^{th}$ element of (M^{-1}), equivalent to $(s, t)^{th}$ element of $(M^{-1})'$.
  (iv) Replace all other elements of A by zero.
  (v) Transpose is the final resulting matrix.
  (vi) The result is G, a generalized inverse of A.
- This factorization is implemented in every professional level statistical package. The Moore-Penrose generalized inverse is produced by :

$$C = [r^{-1}\ 0]\ Q'$$

where, 0 is the transpose of the zero matrix of the R matrix required for conformability.

Miscellaneous Exercise

(A) Theory Questions :

1. Give the definitions of g-inverse and prove their equivalence.
2. Prove or disprove that g-inverse is unique.
3. Find Mp-g inverse of a row vector $[3, 2 - 5]$ and $\begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}$.
4. Define Mp-g inverse. Is this g-inverse is unique ? Justify your answer.

(B) Numerical Problems :

1. Find the generalized inverse of the following matrices.

$$A_1 = \begin{bmatrix} 4 & 1 & 2 & 0 \\ 1 & 1 & 5 & 15 \\ 3 & 1 & 3 & 5 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 2 & 2 & 6 \\ 2 & 3 & 8 \\ 6 & 8 & 22 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 2 & 3 & 1 & -1 \\ 5 & 8 & 0 & 1 \\ 1 & 2 & -2 & 3 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 10 & 7 \\ 2 & 1 & 1 & 6 \end{bmatrix}$$

2. Find the Mp-g inverse of the following matrix C where, $C = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 5 \\ 0 & 1 & -1 \\ 1 & 3 & -1 \end{bmatrix}$.
3. Calculate the Moore Penrose and two other generalized inverses of
   $X = [7 \quad 2 \quad 3]$
4. Prove that $B^- A^-$ is a generalized inverse of AB if and only if A^-ABB^- is idempotent matrix.
5. Prove that Rank (G) = Rank (A) for AGA = A if and only if G is reflexive generalized inverse.
6. When M is idempotent and Z^- is generalized inverse of Z = MAM. Prove that MZ^-M is also generalized inverse of Z.
7. Suppose G is generalized inverse of matrix A then show that G is also generalized inverse of AG if and only if G^2 is a generalized inverse of A.
8. Suppose G is Moore Penrose generalized inverse of symmetric matrix A. Show that G^2 is the Moore Penrose inverse of A^2.
9. If G is generalized inverse of symmetric matrix A then show that G^2 is a generalized inverse of A^2 if GA is symmetric matrix.
10. **Prove that :**
    (a) Rank (AB) = Rank (A) implies $B(AB)^-$ is a generalized inverse of A.
    (b) Rank (AB) = Rank (B) implies $(AB)^- A$ is a generalized inverse of A.

❑❑❑

Chapter 3...

Vector Spaces

Synopsis

3.1 Introduction of Real Vector Space

3.2 Subspaces

3.3 Linear Combination and Spanning Sets

3.4 Linear Dependence and Independence

3.5 Basics and Dimensions

3.6 Row Space, Column Space, Null Space, Rank and Nullity

Objectives

(1) You will be able to understand the concept of different spaces like vector spaces, subspaces, row space, column space and Null space with their basis and dimension.

(2) You will be able to differentiate between independent vectors and set of vectors.

Stefan Banach

Stefan Banach (March 30, 1892 – August 31, 1945) was a Polish mathematician. He is generally considered to have been one of the 20th century's most important and influential mathematicians. Banach was one of the founders of modern functional analysis and one of the original members of the Lwów School of Mathematics. His major work was the 1932 book, *Théorie des opérations linéaires* (Theory of Linear Operations), the first monograph on the general theory of functional analysis.

Some of the notable mathematical concepts named after Banach include Banach spaces, Banach algebras, the Banach–Tarski paradox, the Hahn–Banach theorem, the Banach–Steinhaus theorem, the Banach-Mazur game, the Banach–Alaoglu theorem and the Banach fixed-point theorem.

3.1 Introduction and Real Vector Space

In mathematics we come across many examples of mathematical objects that can be added to each other and multiplied by real numbers. First of all the real numbers themselves are such objects. Other examples are vectors in plane, vectors in space, matrices and the real valued functions. In this chapter, we will study a more theoretical mathematics but will provide a powerful tool for extending our geometrical visualization to a wide variety of important mathematical problems where geometric intuition would not otherwise be available. The vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ can be visualized as arrows, which enables us to draw or form a mental picture to help solve problems. In this chapter we study a general mathematical concept called real vector spaces or real linear spaces and the axioms, which define new kind of vectors, will be based on the properties of vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$. This will help us to solve problems involving our new kind of vectors say matrices or functions.

In chapters 1, 2 and 3 by **scalars** we shall always mean **real numbers** unless and otherwise it is specifically mentioned.

The definition of a real vector space :

Let V be an arbitrary non-empty set of objects on which two operations are defined addition and multiplication by scalars called *sum and scalar multiplication* respectively. If the following axioms are satisfied by all objects u, v, and w in V and for all scalars k and *l*, then V is called a *real vector space* or *real linear space*.

Closure axioms

C_1 : For all u, v in V, u + v is in V.

C_2 : For any scalar k and any object u in V, ku is in V.

Addition axioms

A_1 : u + v = v + u (commutativity of sum)

A_2 : (u + v) + w = u + (v + w) (associativity of sum)

A_3 : There is an object 0 in V, called *zero vector for V,* such that,

u + 0 = u for all u in V.

A_4 : For each u in V, there is an object – u in V, called *negative of u* such that u + (–u) = 0.

Scalar Multiplication axioms

M_1 : k (u + v) = ku + kv

M_2 : (k + *l*) u = ku + *l*u

M_3 : (k*l*) u = k (*l*u)

M_4 : 1u = u.

Note 1 : *Objects of V are called vectors, and hence we may denote elements of vector space by an arrow on the top. For example, for u, v in V, we will write* $\vec{u}$, $\vec{v}$.

Note 2 : *As we are using scalars as real numbers, we call real vector space or real linear space, in this text we will study only real vector spaces, hence will use only* ***vector space*** *for brevity.*

SOLVED EXAMPLES

Example 3.1 : The set of real numbers $\mathbb{R}$ itself is a vector space with addition of any two members in $\mathbb{R}$ as usual addition of real numbers and multiplication by scalars as usual multiplication. The axioms of a vector space clearly follow from the properties of addition and multiplication of real numbers.

Example 3.2 : Let $V = \mathbb{R}^2$ be the set of all ordered pairs of real numbers, that is,

$\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$. Define for any $\vec{u} = (x, y)$ and $\vec{v} = (x', y')$ in $\mathbb{R}^2$

(i) Equality : $\vec{u} = \vec{v}$ if and only if $x = x'$ and $y = y'$.

(ii) Sum : $\vec{u} + \vec{v} = (x, y) + (x', y')$
$= (x + x', y + y')$

(iii) Scalar Multiplication : For any scalar k

$k\vec{u} = k(x, y) = (kx, ky)$.

Show that $\mathbb{R}^2$ is a vector space.

Solution : Let us show that all the vector space axioms are satisfied by these two operations in $\mathbb{R}^2$.

The closure axioms clearly hold good as $x + x'$, $y + y'$ and kx, ky are real numbers, hence, $\vec{u} + \vec{v}$ and $k\vec{u}$ are again ordered pairs of real numbers and thus belong to $\mathbb{R}^2$.

Let $\vec{u} = (x, y)$; $\vec{v} = (x', y')$ and $\vec{w} = (x'', y'')$ be any vectors in $\mathbb{R}^2$.

A_1 : $\vec{u} + \vec{v} = (x, y) + (x', y')$

$= (x + x', y + y')$, by definition of sum in $\mathbb{R}^2$.

$= (x' + x, y' + y)$ $\because$ sum in $\mathbb{R}$ is commutative.

$= (x', y') + (x, y)$, again by definition of sum in $\mathbb{R}^2$.

$= v + u$.

$\therefore \quad \vec{u} + \vec{v} = \vec{v} + \vec{u}$, for all u and v in $\mathbb{R}^2$.

A_2 : $(\vec{u} + \vec{v}) + \vec{w} = [(x, y) + (x', y')] + (x'', y'')$,

$= (x + x', y + y') + (x'', y'')$ by sum in $\mathbb{R}^2$

$= ((x + x') + x'', (y + y') + y'')$, by sum in $\mathbb{R}^2$.

$= (x + (x' + x''), y + (y' + y''))$, by associativity of sum in $\mathbb{R}$.

$= (x, y) + (x' + x'', y' + y'')$, by sum in $\mathbb{R}^2$.

$= (x, y) + [(x', y') + (x'' + y'')]$, by sum in $\mathbb{R}^2$.

$= \vec{u} + (\vec{v} + \vec{w})$

$\therefore \quad (\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$,

A_3 : As $0 \in \mathbb{R}$, (0, 0) is in $\mathbb{R}^2$, and we have,

$$\begin{aligned}
\vec{u} + (0, 0) &= (x, y) + (0, 0) \\
&= (x + 0, y + 0), && \text{by sum in } \mathbb{R}^2. \\
&= (x, y) \\
&= \vec{u}.
\end{aligned}$$

Thus, (0, 0) is a zero vector in $\mathbb{R}^2$, we will denote it by $\vec{0}$.

A_4 : For $\vec{u} = (x, y)$ in $\mathbb{R}^2$, we have $-\vec{u} = (-x, -y)$ in $\mathbb{R}^2$ such that,

$$\begin{aligned}
\vec{u} + (-\vec{u}) &= (x, y) + (-x, -y) \\
&= (x + (-x), (y + (-y)), && \text{by sum in } \mathbb{R}^2. \\
&= (0, 0) && \because x + (-x) = 0 = y + (-y) \text{ in } \mathbb{R}
\end{aligned}$$

Thus, for each $\vec{u}$ in $\mathbb{R}^2$, there is negative of $\vec{u}$ in $\mathbb{R}^2$.

M_1 : Now, let k and l be any scalars. Consider

$$\begin{aligned}
k(\vec{u} + \vec{v}) &= k(x + x', y + y'), && \text{by sum in } \mathbb{R}^2 \\
&= (k(x + x'), k(y + y')), && \text{by scalar multiplication in } \mathbb{R}^2. \\
&= (kx + kx', ky + ky') && \because \text{multiplication distributes over '+' in } \mathbb{R} \\
&= (kx, ky) + (kx', ky'), && \text{by sum in } \mathbb{R}^2 \\
&= k(x, y) + k(x', y'), && \text{by scalar multiplication in } \mathbb{R}^2. \\
&= k\vec{u} + k\vec{v}
\end{aligned}$$

Thus, $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$.

M_2 : Consider,

$$\begin{aligned}
(k + l)\vec{u} &= (k + l)(x, y) \\
&= ((k + l)x, (k + l)y), && \text{by scalar multiplication in } \mathbb{R}^2. \\
&= (kx + lx, ky + ly), && \because \text{multiplication distributes over + in } \mathbb{R} \\
&= (kx, ky) + (lx, ly), && \text{by sum in } \mathbb{R}^2. \\
&= k(x, y) + l(x, y), && \text{by scalar multiplication in } \mathbb{R}^2 \\
&= k\vec{u} + l\vec{u}.
\end{aligned}$$

Thus, $(k + l)\vec{u} = k\vec{u} + l\vec{u}$.

M_3 : Consider

$$(kl)\,\vec{u} = (kl)\,(x, y)$$
$$= ((kl)x, (kl)\,y), \quad \text{by scalar multiplication in } \mathbb{R}^2$$
$$= (k\,(lx)\,,\, k\,(ly)) \quad \because \text{Multiplication is associative in } \mathbb{R}^2$$
$$= k\,[l(x, y)], \quad \text{by scalar multiplication}$$
$$= k\,(l\,\vec{u}\,).$$

Thus, $(kl)\,\vec{u} = k(l\,\vec{u}\,)$

M_4 : $1\,\vec{u} = 1\,(x, y)$
$= (1x, 1y)$
$= (x, y)$
$= \vec{u}$

Thus, $\mathbb{R}^2$ is a vector space.

Example 3.3 : Let $V = \mathbb{R}^n$, the set of all ordered n-tupples of real numbers, that is,

$\mathbb{R}^n = \{(u_1, u_2, \ldots\ldots, u_n)/\ u_1, u_2, \ldots\ldots, u_n \text{ are real numbers}\}$

For any $\vec{u} = (u_1, u_2\,, \ldots\ldots, u_n)$, $\vec{v} = (v_1, v_2, \ldots\ldots, v_n)$ in $\mathbb{R}^n$ and any scalar k, define.

(i) Equality : $\vec{u} = \vec{v}$ if and only if $u_1 = v_1$, $u_2 = v_2$, ……, $u_n = v_n$.

(ii) Sum : $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \ldots\ldots, u_n + v_n)$

and **(iii) Scalar multiplication :** $k\,\vec{u} = (ku_1, ku_2, \ldots\ldots\ ku_n)$.

Then $\mathbb{R}^n$ is a vector space can be verified analogously as in Example 3.2.

This is finite generalisation of example (3.2). One can verify all axioms of vector space and $\mathbb{R}^n$ is a vector space.

Note : The addition and scalar multiplication defined as above in $\mathbb{R}^2$ and $\mathbb{R}^n$ is called componentwise addition and componentwise scalar multiplication, which is also called *usual addition and scalar multiplication.*.

Example 3.4 : Let $V = M_{2\times 2}(\mathbb{R})$, the set of all 2×2 matrices with real entries.

For any $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

and $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ in V,

we know that,

(i) Equality : A = B iff a = e, b = f, c = g, d = h.

(ii) Sum : $A + B = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$

and (iii) Scalar multiplication

$$kA = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}.$$ Show that V is a vector space.

Solution : To see that V satisfies all the axioms of vector space, first closure axioms and follow from definition of addition and scalar multiplication.

Axiom A_1 follows, because

$$\begin{aligned} A + B &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} \\ &= \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} \\ &= \begin{bmatrix} e+a & f+b \\ g+c & h+d \end{bmatrix} && (\because \text{ a, b, c, d, e, f, g, h are real number}) \\ &= \begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix}, && \text{by definition of addition in V.} \\ &= B + A \end{aligned}$$

Axioms A_2, M_1, M_2 and M_3 follow by using the properties of real numbers and definitions of addition and scalar multiplication in V. For axiom A_3, we have to find $0 \in V$ such that $A + O = O + A = A$, for all $A \in V$. This is done by defining O to be

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

which clearly belongs to V and

$$O + A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = A$$

To prove axiom A_4 given any $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in V, we have – A in V defined as

$$-A = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$$

With this definition, we have

$$A + (-A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

Lastly, to see axiom M_4 is satisfied, because

$$1 \cdot A = 1 \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 \cdot a & 1 \cdot b \\ 1 \cdot c & 1 \cdot d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = A$$

Thus, $V = M_{2\times 2}(\mathbb{R})$ is a vector space.

Example 3.5 : Let $V = M_{m \times n}(\mathbb{R})$ be the set of all $m \times n$ matrices with real entries. Then V is a vector space with operations of addition of matrices and scalar multiplication.

Solution : The arguments in example (4) can be used to show that V is a vector space. The $m \times n$ zero matrix is the zero vector in V, and if A is any $m \times n$ matrix in V, then $-A$ is the negative of the matrix A.

Note : If $m = n$, then we denote by $M(n, \mathbb{R})$, the set of all $n \times n$ matrices with real entries. Clearly $M(n, \mathbb{R})$ is also a vector space with the operations of addition and scalar multiplication of matrices.

Example 3.6 : Let S_n denote the set of all $n \times n$ symmetric real matrices. (Clearly $S_n \subset M(n, \mathbb{R})$). Then under the operations of matrix addition and scalar multiplication, S_n is a vector space.

Example 3.7 : If A_n denotes set of all skew-symmetric matrices of order $n \times n$, then A_n is a vector space under usual addition and scalar multiplication of matrices. (Clearly A_n is also subset of $M(n, \mathbb{R})$.

Example 3.8 : Let V and W be vector spaces. Let us form the Cartesian product

$$V \times W = \{(v, w) | v \in V \text{ and } w \in W\}.$$

Define addition and scalar multiplication on $V \times W$ as

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2), (v_1, w_1), (v_2, w_2) \in V \times W$$

and $$\alpha (v, w) = (\alpha v, \alpha w), \alpha \in \mathbb{R} \text{ and } (v, w) \in V \times W.$$

Then $V \times W$ is a vector space. This vector space is usually denoted by $V \oplus W$ and called the *direct sum of V and W.*

Note : We can extend the concept of direct sum in example (8) to define direct sum $V_1 \oplus V_2 \oplus \ldots \oplus V_n$ of n vector spaces V_i, $1 \leq i \leq n$. Then it is clear that

$$\mathbb{R}^n = \underbrace{\mathbb{R} \oplus \mathbb{R} \oplus \ldots \oplus \mathbb{R}}_{\text{n-times}}$$

Let V be a vector space. On $V \times V$, define + and $\cdot$ as follows :

(a) $$(v_1, w_1) + (v_2, w_2) = (v_1 + w_2, w_1 + v_2)$$
$$\alpha (v, w) = (\alpha v, \alpha w), \alpha \in \mathbb{R}, (v, w) \in V \times V$$

(b) $$(v_1, w_1) + (v_2, w_2) = (v_1 + w_1, v_2 + w_2)$$
$$\alpha (v, w) = (\alpha v, \alpha w).$$

Reasons : In both the cases (a) and (b), $V \times V$ is not a vector space. In case of (a) the axiom of addition commutative (A_1) is not satisfied. In case (b) also axiom (A_1) addition commutative is not satisfied.

Example 3.9 : Let $V = F([a, b], \mathbb{R})$ be the set of all real valued functions defined on a closed interval [a, b], $a < b$. For any f and g in V and for any scalar k, define,

(i) **Equality :** $f = g$ if and only if $f(x) = g(x)$ for all x in [a, b].

(ii) **Addition :** $f + g : [a, b] \to \mathbb{R}$ given by $(f + g)(x) = f(x) + g(x)$, $x \in [a, b]$.

and (iii) **Scalar multiplication :** $kf : [a, b] \to \mathbb{R}$ given by $(kf)(x) = k f(x)$, $x \in [a, b]$.

Show that V is a vector space.

Solution : Let us verify that V is a vector space under the sum and scalar multiplication defined as above.

Firstly, C_1 and C_2 follow clearly as f + g and kf are again real valued functions defined on [a, b].

A_1 : For any x in [a, b],

$$\begin{aligned}(f + g)(x) &= f(x) + g(x), && \text{by sum in V.}\\ &= g(x) + f(x), && \because f(x) \text{ and } g(x) \text{ are real numbers.}\\ &= (g + f)(x), && \text{by sum in V.}\end{aligned}$$

$\therefore$ $(f + g)(x) = (g + f)(x)$, for all x in [a, b]. Hence, by equality in V, we have

$$f + g = g + f$$

and A_1 follows.

A_2 : For any $x \in [a, b]$, we have,

$$\begin{aligned}[(f + g) + h](x) &= (f + g)(x) + h(x), && \text{by definition of sum in V.}\\ &= [f(x) + g(x)] + h(x) && \text{by definition of sum in V.}\\ &= f(x) + [g(x) + h(x)], && \text{by associativity of + in } \mathbb{R}.\\ &= f(x) + (g + h)(x), && \text{by definition of + in V.}\\ &= [f + (g + h)](x), && \text{by definition of + in V.}\end{aligned}$$

$\therefore$ $(f + g) + h = f + (g + h)$ by definition of equality

A_3 : For each $x \in [a, b]$, we define a function $\overline{0} : [a, b] \to \mathbb{R}$ by $\overline{0}(x) = 0$, then for any $f \in V$,

$$\begin{aligned}(f + \overline{0})(x) &= f(x) + \overline{0}(x) && x \in [a, b].\\ &= f(x) + 0 && x \in [a, b].\\ &= f(x) && x \in [a, b].\end{aligned}$$

$\therefore$ $f + \overline{0} = f$

$\overline{0}$ is zero vector in V.

A_4 : For any $f \in V$, we define,

$$\begin{aligned}-f : [a, b] &\to \mathbb{R} \text{ as}\\ (-f)(x) &= -f(x), \text{ then we have } -f \in V \text{ and}\\ [f + (-f)](x) &= f(x) + (-f(x)), && \text{by definition of '+' in V.}\\ &= f(x) - f(x).\\ &= 0\end{aligned}$$

Thus, $[f + (-f)](x) = 0$ for all x.

$$= \overline{0}(x) \qquad \text{for all x.}$$

$\therefore$ $f + (-f) = \overline{0}$

M_1 : For any $x \in [a, b]$ and for any f, g in V and k any scalar consider,

$[k (f + g)] (x) = k [(f + g) (x)]$, by definition of scalar multiplication

$= k (f (x) + g (x))$ by definition of sum in V.

$= k f (x) + k g (x)$, by distributive law for $\mathbb{R}$.

$= (kf) (x) + (kg) (x)$, by definition of scalar multiplication

$= (kf + kg) (x)$, by definition of sum in V.

$\therefore \quad [k (f + g)] (x) = (kf + kg) (x)$, for all $x \in [a, b]$.

$\therefore \quad k (f + g) = kf + kg$, by definition of equality.

M_2 : For any scalars k and l and $f \in V$ consider,

$[(k + l)f] (x) = (k + l) f (x)$, by definition of scalar multiplication

$= k f (x) + l f(x)$, by distributive law in $\mathbb{R}$.

$= (kf (x) + (l f) (x)$, by definition of scalar multiplication

$= (kf) + l f) (x)$, by definition of sum.

Thus,$[(k + l) f] (x) = (kf + l f) (x)$, for all $x \in [a, b]$.

$\therefore \quad (k + l)f = kf + lf.$

M_3 : For any scalars k and l and any $f \in V$, consider,

$[(kl) f] (x) = (kl) f (x)$, by definition of scalar multiplication,

$= k (lf (x))$, by associativity of multiplication in $\mathbb{R}$.

$= k [(lf) (x)]$, by definition of scalar multiplication.

$= [k (l f)] (x)$, by definition of scalar multiplication.

Thus, $[(kl) f] (x) = [k (lf)] (x)$, for all $x \in [a, b]$.

$\therefore \quad (kl) f = k (lf).$

M_4 : For any $f \in V$, and any $x \in [a, b]$,

$(1f) (x) = 1 f (x)$, by definition of scalar multiplication.

$= f(x).$

Thus, $(1f) (x) = f (x)$, for all $x \in [a, b]$.

$\therefore \quad 1f = f$

Note :

(1) The above vector space is sometimes called a function space.

(2) We do not use arrow on the top of elements of vector space of matrices and vector space of functions.

Example 3.10 : Let V = C [a, b], be the set of all real valued *continuous* functions defined on [a, b]. We define the addition and scalar multiplication as in example (3.9). It is easy to see that f + g and αf are in V if $f, g \in V$, because we know from analysis that sum of two continuous functions is continuous, also a scalar multiple of a continuous function is continuous. The rest of the axioms follow similarly as in example (3.9).

Note : C [a, b] is a subset F ([a, b], $\mathbb{R}$).

Example 3.11 : The set V = D ([a, b]) of differentiable functions on [a, b] is a vector space under the addition and scalar multiplication of functions as in example (3.9). Here, D ([a, b]) is a subset of F ([a, b], $\mathbb{R}$).

Example 3.12 : The set V = R ([a, b]), of all Riemann integrable functions on [a, b], which is also a subset of F([a, b], $\mathbb{R}$) is a vector space under addition and scalar multiplication as defined in example (3.9).

Example 3.13 : Let $V = P = \left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in \mathbb{R},\ n \in \mathbb{N} \right\}$, be the set of all polynomials in one variable with real coefficients. The addition and scalar multiplication are defined as

$$\left(\sum_{i=0}^{m} a_i x^i \right) + \left(\sum_{j=0}^{n} b_j x^j \right) = \sum_{r} (a_r + b_r)\, x^r,$$

where, $a_r = 0$ if $r > m$ and $b_r = 0$ if $r > n$. Also,

$$\alpha \left(\sum_{i=0}^{n} a_i x^i \right) = \sum_{i=0}^{n} (\alpha a_i)\, x^i.$$

Then V is a vector space over $\mathbb{R}$. This vector space is called a polynomial vector space.

Example 3.14 : Let $V = P_n = \left\{ \sum_{i=0}^{n} a_i x^i \mid a_i \in \mathbb{R} \right\}$, be the set of all polynomials of degree $\leq n$ with real coefficients. Define operations as in P, then V is a vector space over $\mathbb{R}$. Note that P_n is a subset of P.

Note 1. : Let V denote the set of all polynomials exactly of degree n. Then V is **not** vector space under the operations as in P; because 0 element cannot be in V as it is not a polynomial of degree n. Also one can see easily that if $\alpha = 0$ and P(x) is any polynomial in V, the $\alpha P(x)$ is not in V.

Note 2 : If $x \in [a, b]$, then P and P_n are both subsets of F ([a, b], $\mathbb{R}$).

Uptill now, we have seen the set V of objects, which were already known to us and the sum and scalar multiplication that we have defined on V were also known to us from our prior knowledge of mathematics. Now, let us see one example in which the given set of objects is known to us and we can define different operations of addition and scalar multiplication, which are unusual.

Example 3.15 : Let $V = \mathbb{R}^+$ be the set of all positive reals. Define addition of any two members $\vec{x}$ and $\vec{y}$ to be the usual multiplication of numbers, that is, $\vec{x} + \vec{y} = x \cdot y$. Define scalar multiplication by a scalar k to any $\vec{x} \in \mathbb{R}^+$ to be x^k, that is, $k\vec{x} = x^k$. Then show that V is a vector space.

Solution : Let us verify that V is a vector space.

C_1 : Since the product of two positive real numbers is again positive real number, for any $\vec{x}, \vec{y} \in \mathbb{R}^+$, $x + y \in \mathbb{R}^+$.

C_2 : For any real number k and $\vec{x} \in \mathbb{R}^+$, x^k is also a positive number, hence $k\vec{x} \in \mathbb{R}^+$.

For any $\vec{x}$, $\vec{y}$ and $\vec{z}$ in $\mathbb{R}^+$, we have,

A_1 : $\vec{x} + \vec{y} = x \cdot y$ by definition of '+' in $\mathbb{R}^+$

$= y \cdot x$ $\because$ multiplication in $\mathbb{R}^+$ is commutative.

$= \vec{y} + \vec{x}$, again by definition of '+' in $\mathbb{R}^+$

$\therefore$ $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.

A_2 : $(\vec{x} + \vec{y}) + \vec{z} = (x \cdot y) + \vec{z}$, by definition of '+' in $\mathbb{R}^+$

$= (x \cdot y) \cdot z$ by definition of '+' in $\mathbb{R}^+$

$= x \cdot (y \cdot z)$ by associativity of multiplication in $\mathbb{R}^+$

$= \vec{x} + (y \cdot z)$, by definition of '+' in $\mathbb{R}^+$

$= \vec{x} + (\vec{y} + \vec{z})$, by definition of '+' in $\mathbb{R}^+$

Thus, $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$.

A_3 : $1 \in \mathbb{R}^+$, such that for any $\vec{x} \in \mathbb{R}^+$

$$\vec{x} + \vec{1} = x \cdot 1 = \vec{x}$$

and $$\vec{1} + \vec{x} = 1 \cdot x = \vec{x},$$

Thus, $\vec{1}$ is zero element in $\mathbb{R}^+$.

Note : From this example we see that zero element of a vector space may not be usual zero element, it is zero element with respect to the operation defined in the set.

A_4 : For any $\vec{x} \in \mathbb{R}^+$, $\left(\frac{\vec{1}}{x}\right) \in \mathbb{R}^+$, $\because x > 0$.

s.t. $\vec{x} + \left(\frac{\vec{1}}{x}\right) = x \cdot \frac{1}{x}$ by definition of '+' in $\mathbb{R}^+$

$= \vec{1}$, the zero element.

Thus, each element in $\mathbb{R}^+$ has negative element in $\mathbb{R}^+$.

M_1 : For any $\vec{x}$, $\vec{y} \in \mathbb{R}^+$ and any scalar k, consider,

$$\begin{aligned} k(\vec{x} + \vec{y}) &= k(x \cdot y), && \text{by definition of } + \text{ in } \mathbb{R}^+ \\ &= (x \cdot y)^k, && \text{by definition of scalar multiplication.} \\ &= x^k \cdot y^k, && \text{by laws of indices.} \\ &= x^k + y^k, && \text{by definition of sum in } \mathbb{R}^+. \\ &= k\vec{x} + k\vec{y} && \text{by definition of scalar multiplication.} \end{aligned}$$

Thus, $k(\vec{x} + \vec{y}) = k\vec{x} + k\vec{y}$.

M_2 : For any $\vec{x} \in \mathbb{R}^+$ and for any scalars k and l, consider,

$$\begin{aligned} (k + l)\vec{x} &= x^{(k+l)}, && \text{by definition of scalar multiplication.} \\ &= x^k . x^l, && \text{by laws of indices.} \\ &= x^k + x^l, && \text{by definition of '+' in } \mathbb{R}^+ \\ &= k\vec{x} + l\vec{x}, && \text{by definition of scalar multiplication} \end{aligned}$$

Thus, $(k + l)\vec{x} = k\vec{x} + l\vec{x}$.

M_3 : For any $\vec{x} \in \mathbb{R}^+$ and any scalar k and l, consider

$$\begin{aligned} (kl)\vec{x} &= x^{kl}, && \text{by definition of scalar multiplication} \\ &= (x^l)^k, && \text{by laws of indices.} \\ &= k(x^l), && \text{by definition of scalar multiplication} \\ &= k(l\vec{x}), && \text{again by definition of scalar multiplication} \end{aligned}$$

thus, $(kl)x = k(l\vec{x})$.

M_4 : For any $\vec{x} \in \mathbb{R}^+$, we have

$$\begin{aligned} 1 \cdot \vec{x} &= x^1 && \text{by definition of scalar multiplication} \\ &= \vec{x}, && \text{since } x^1 = x \end{aligned}$$

Thus, $1 \cdot \vec{x} = \vec{x}$.

Thus, $\mathbb{R}^+$ is a vector space with 1 as zero element and negative of any $\vec{x}$ as $\left(\frac{\vec{1}}{x}\right)$.

Note : We have used arrow on the top of elements of $\mathbb{R}^+$ as a vector space, while we have not used arrow on the top, when we considered them as real numbers.

Example 3.16 : Let $V = M_{2\times 2}(\mathbb{R})$

For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$, define,

$$A + B = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}.$$

For any scalar k, define,

$$kA = \begin{bmatrix} ka & 0 \\ 0 & kd \end{bmatrix}$$

Solution : With the operations defined as above, does not satisfy all the axioms of a vector space, though closure axioms C_1 and C_2 and axioms of addition A_1, A_2, A_3 and A_4 are satisfied by V. (Check !). However, axioms regarding scalar multiplication are not all satisfied. As a case, M_4 is not satisfied; because for any

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$1A = \begin{bmatrix} 1a & 0 \\ 0 & 1d \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \text{ which is not A itself.}$$

That is, in this case, $1 \cdot A \neq A$.

$\therefore$ V is not a vector space under these operations.

Elementary Consequences of Axioms of Vector Space :

Now, let us see some consequences of axioms of a vector space.

Theorem 1 : Let V be a vector space. Then

(1) There is only one zero in V.

(2) Each vector in V has just one negative in V.

Proof : (1) Suppose that $\vec{0}$ and $\vec{0'}$ are both zero objects in V. We have,

$$\vec{0} = \vec{0} + \vec{0'}, \qquad \text{since } \vec{0'} \text{ is a zero in V.}$$

Also, $$\vec{0'} = \vec{0} + \vec{0'} \qquad \text{since } \vec{0} \text{ is a zero in V.}$$

We get, $$\vec{0'} = \vec{0}$$

Note : As we have proved that there is unique zero vector in any vector space, we may speak of *the zero* vector in V.

(2) Let $\vec{u}$ be a vector in V. Suppose that $\vec{y}$ and $\vec{z}$ are both negatives of $\vec{x}$ in V. We will show that $\vec{y} = \vec{z}$. We have,

$$\vec{y} = \vec{0} + \vec{y}, \qquad \because \text{'0' is zero in V}$$

$$= (\vec{z} + \vec{x}) + \vec{y}, \qquad \because \vec{z} \text{ is negative of } \vec{x},\ \vec{z} + \vec{x} = \vec{0}.$$

$= \vec{z} + (\vec{x} + \vec{y})$, by axiom A_2.

$= \vec{z} + \vec{0}$, $\because \vec{y}$ is negative of $\vec{x}$, $\vec{x} + \vec{y} = \vec{0}$.

$= \vec{z}$

Thus, $\vec{y} = \vec{z}$.

Note :

(1) We may speak now, *the negative* of a vector $\vec{u}$ in V, and we shall denote the negative of $\vec{u}$ by $-\vec{u}$.

(2) With above convention, we may define subtraction in $\vec{V}$ as : For $\vec{u}$ and $\vec{v}$ in V,

$$\vec{u} - \vec{v} = \vec{u} + (-\vec{v}).$$

Theorem 2 : Let V be a vector space, $\vec{u}$ a vector in V and k be any scalar. Then,

(i) $0 \cdot \vec{u} = \vec{0}$

(ii) $k \cdot \vec{0} = \vec{0}$

(iii) $(-1) \cdot \vec{u} = -\vec{u}$

(iv) $k \cdot \vec{u} = \vec{0}$ if and only if either $k = 0$ or $\vec{u} = \vec{0}$.

Proof : (i) We have,

$0 + 0 = 0$

$\therefore (0 + 0)\vec{u} = 0 \cdot \vec{u}$

$\Rightarrow 0.\vec{u} + 0.\vec{u} = 0.\vec{u}$, by axiom A_4.

$\Rightarrow [0.\vec{u} + 0.\vec{u}] + (-0.\vec{u}) = 0.\vec{u} + (-0.\vec{u})$, by adding negative of $0.\vec{u}$ on both sides,

$\Rightarrow 0.\vec{u} + [0.\vec{u} + (-0).\vec{u}] = 0.\vec{u} + (-0.\vec{u})$, by axiom A_2.

$\Rightarrow 0.\vec{u} + \vec{0} = \vec{0}$, by axiom A_4.

$\Rightarrow 0.\vec{u} = \vec{0}$, by axiom A_3.

(ii) Again for $\vec{0} \in V$, we have,

$k . (\vec{0} + \vec{0}) = k \cdot \vec{0}$, by axiom A_3.

$\Rightarrow k.\vec{0} + k.\vec{0} = k.\vec{0}$ by axiom M_1.

Now, adding negative of $k\cdot\vec{0}$ on both sides we get in similar manner as in (i), that $k.\vec{0} = \vec{0}$.

(iii) To show that $(-1)\vec{u} = -\vec{u}$, means we have to show that $(-1)\vec{u}$ is negative of $\vec{u}$. Therefore, consider,

$$\vec{u} + (-1)\vec{u} = 1.\vec{u} + (-1)\vec{u}, \qquad \text{by axiom } M_4.$$

$$= (1 + (-1)).\vec{u}, \qquad \text{by axiom } M_2.$$

$$= 0.\vec{u}, \qquad \because 1 + (-1) = 0.$$

$$= \vec{0}, \qquad \text{by (i) above.}$$

$$\therefore \quad \vec{u} + (-1).\vec{u} = \vec{0}$$

implies that $(-1)\vec{u}$ is a negative of $\vec{u}$, that is $(-1).\vec{u} = -\vec{u}$.

(iv) If either $k = 0$ or $\vec{u} = \vec{0}$, the case turns out to be (i) or (ii) above, and $k.\vec{u} = \vec{0}$ in either case follows.

Conversely, suppose $k.\vec{u} = \vec{0}$. If $k = 0$, then there is nothing to be proved. So, suppose $k \neq 0$, then k being a scalar, there is $\frac{1}{k}$ such that $k.\frac{1}{k} = 1$.

Now, multiplying both sides by $\frac{1}{k}$, we get,

$$\frac{1}{k}(k\vec{u}) = \frac{1}{k}\vec{0}$$

$$\Rightarrow \quad \frac{1}{k}(k\vec{u}) = \vec{0}, \qquad \text{by (ii) above.}$$

$$\Rightarrow \quad \left(\frac{1}{k}k\right)\vec{u} = \vec{0}, \qquad \text{by axiom } M_3.$$

$$\Rightarrow \quad 1\,\vec{u} = \vec{0}$$

$$\Rightarrow \quad \vec{u} = \vec{0}, \qquad \text{by axiom } M_4.$$

Exercise 3.1

In Exercises 1 – 12 set V of objects is given together with operations of addition and scalar multiplication. Determine which sets are vector spaces under given operations. For those that are not, state the axioms that fail to hold.

1. $V = \mathbb{R}^2$, set of ordered pairs of real numbers with operations
   $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $k(x_1, y_1) = (kx_1, y_1)$.

2. $V = \mathbb{R}^2$, with operations $(x_1, y_1) + (x_2, y_2) = (x_1, y_1 + y_2)$ and $k(x_1, y_1) = (kx_1, ky_1)$.
3. $V = \mathbb{R}^3$ with the operations
   $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ and $k(x_1, y_1, y) = (kx_1, 0, 0)$.
4. $V = \mathbb{R}^3$ with the operations
   $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$ and $k(x_1, y_1, z_1) = (2kx_1, 2ky_1, 2kz_1)$.
5. $V = \mathbb{R}^3$, with operations
   $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2 - 1, y_1 + y_2 - 1, z_1 + z_2 - 1)$ and $k(x_1, y_1, z_1) = (kx_1, ky_1, kz_1)$.
6. $V = \{(x, y, 0) | x, y \in \mathbb{R}\}$ with the usual operations of addition and scalar multiplication.
7. $V = \{(x, y, z) \in \mathbb{R}^3 \mid \text{either } x = 0 \text{ or } y = 0\}$. With usual operations of addition and scalar multiplication.
8. $V = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, z) = k(3, 2, 1), \text{ with } k \in \mathbb{R}\}$.
9. $V = \{(x, y, z) \in \mathbb{R}^3 \mid y = 5x\}$.
10. $V = \mathbb{R}^2$ with operations $(x_1, y_1) + (x_2, y_2) = (|x_1 + x_2|, |x_2 + y_2|)$ and $k(x_1, y_1) = (kx_1, ky_2)$.
11. $V = \mathbb{R}^2$ with operations $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, 0)$ and $k(x_1, y_1) = (kx_1, ky_1)$.
12. $V = \mathbb{R}^2$ with operations $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and
    $k(x_1, y_1) = (|k|x_1, |k|y_1)$.

In exercise 13 through 20. the set V of objects in the set of functions which are real valued with domain R, with addition and scalar multiplication as defined in Example 3.9. Determine which sets are vector spaces. For those that are not, list all axioms that fail to hold.

13. $V = \{f \mid f(0) = f(2)\}$.
14. $V = \{f \mid f \text{ is an even function}\}$.
15. $V = \{f \mid f \text{ is an odd function}\}$.
16. $V \{f \mid f\left(\frac{1}{2}\right) \geq 0\}$.
17. $V = \{f \mid f\left(\frac{1}{2}\right) = 0.\}$
18. $V = \{f \mid f \text{ is bounded function}\}$.
19. $V = \{f \mid f \text{ is polynomial with constant term zero}\}$.
20. $V = \{f \mid f\left(\frac{1}{2}\right) = \frac{1}{2} f(0)\}$.

Through exercises 21 – 25, the set V of objects is a subset of set of 2×2 matrices with real entries. Determine which sets are vector spaces under the indicated operations. For those, which are not, give reason/s.

21. If V is the set of all 2×2 matrices of the form $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ with matrix addition and scalar multiplication.
22. If V is set of all 2×2 matrices of the form $\begin{bmatrix} 1 & a \\ b & 1 \end{bmatrix}$ with usual matrix addition and scalar multiplication.
23. The set V of all matrices of the form $\begin{bmatrix} a+b & a \\ b & a+b \end{bmatrix}$ with matrix addition and scalar multiplication.
24. The set V of all matrices of the form $\begin{bmatrix} a & b \\ c & 0 \end{bmatrix}$ with matrix addition and scalar multiplication.
25. The set V of all 2×2 matrices with matrix addition and for a scalar k and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, scalar multiplication defined by $kA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Hints and Answers (3.1)

1. Not a vector space. Axiom M_2 fails.
2. Not a vector space. Axioms A_1, A_3, A_4, M_2 fail.
3. Not a vector space. Axiom M_4 fails.
4. Not a vector space. Axiom M_4 fails.
5. Not a vector space. Axiom M_1 and M_2 fail.
6. Yes, it is a vector space.
7. Not, closure axiom c_1 fails.
8. Yes, it is a vector space.
9. Yes, it is a vector space.
10. No, axioms A_2, A_3, A_4, M_1 and M_2 fail.
11. No, axioms A_2, A_3 and A_4 fail.
12. No axiom M_2 fails.
13. Yes, it is a vector space.
14. Yes, it is a vector space.
15. Yes, it is a vector space.
16. No, closure axiom c_2 fails.
17. Yes, it is a vector space.
18. Yes, it is a vector space.
19. Yes, it is a vector space.
20. Yes, it is a vector space.
21. Yes, it is a vector space.
22. No, axiom A_3 fails.
23. Yes, it is a vector space.
24. Yes, it is a vector space.
25. No, axiom M_4 fails.

3.2 Subspaces

Definition : A subset W of a vector space V is called a subspace of V if W itself is a vector space under the induced operations of addition and scalar multiplication defined on V.

Note : Given a subset $\vec{W}$ of a vector space V to check that W is a subspace of V, one need not verify the ten vector space axioms. Since V is already known to be a vector space, then certain axioms need not be verified for W, because they are inherited from V. For example, the axiom A_2, which states that for all $\vec{u}$, $\vec{v}$, $\vec{w}$, $\in$ V, $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$, which will also hold for all $\vec{u}$, $\vec{v}$, $\vec{w}$ in W. Infact, the axioms A_1, A_2, M_1, M_2, M_3 and M_4 need not be verified for W. Thus, we have,

Theorem 3 : Let W be a subset of a vector space V. Then W is subspace of V if and only if the following conditions hold

(i) $W \neq \phi$

(ii) For any $\vec{u}$ and $\vec{v}$ in W, $\vec{u} + \vec{v}$ is in W, and,

(iii) For any $\vec{u}$ in W, for any scalar k, $k\vec{u}$ is also in W.

Proof : If W is a subspace of V, it is a vector space, then all the vector space axioms are satisfied by W and it is non-empty also, so the conditions (i), (ii) and (iii) hold for W.

Conversely, suppose W is a subset of V which satisfies the given three conditions, we have to show that W is a subspace of V, that is, we need to show that W satisfies all the axioms for vector space. However, the axioms A_1, A_2, M_1, M_2, M_3 and M_4 are inherited from V, we need to verify the axioms C_1, C_2 and A_3, A_4, but C_1 and C_2 are nothing but the given conditions (ii) and (iii), hence, we need only to verify A_3 and A_4.

Let $\vec{u}$ be any vector in W, by condition (iii), $k\vec{u}$ is in W for all scalars k, hence setting k = 0, we have $0.\vec{u} = \vec{0} \in$ W ; that is $\vec{0} \in$ W. Again by setting k = –1, we get $(-1)\vec{u} \in$ W, but $(-1)\vec{u} = -\vec{u}$, that is, $-\vec{u} \in$ W. Thus, zero element is in W and each element in W has negative in W, so that W satisfies A_3 and A_4 and hence all the axioms for vector space. Therefore, W is a subspace of V.

Note :

(1) We can restate the theorem 3 as : *'A non-empty subset W of a vector space V is a subspace of V if and only if it is closed under addition and closed under scalar multiplication.'*

(2) The theorem 3 has alternative statement :

'A non-empty subset W of a vector space V is a subspace of V if and only if, for any $\vec{w_1}$ *and* $\vec{w_2}$ *in W and for any scalars* α *and* β, $\alpha\vec{w_1} + \beta\vec{w_2}$ *is in W.'*

Problem : Establish the equivalence of the statement of theorem 3 and the statement in note (2) above.

Result : Let V be a vector space. If W_1 and W_2 are subspaces of V, then the intersection $W_1 \cap W_2$ is also a subspace of V. What about the union $W_1 \cup W_2$?

Proof : Let $\vec{x}$ and $\vec{y}$ be any two vectors in $W_1 \cap W_2$. Then $\vec{x}$, $\vec{y} \in W_1$ and $\vec{x}$, $\vec{y} \in W_2$. Since W_1 and W_2 are subspaces $\vec{x} + \vec{y} \in W_1$ and $\vec{x} + \vec{y} \in W_2$. That is $\vec{x} + \vec{y} \in W_1 \cap W_2$. Also, for any scalar k, W_1 and W_2 being subspaces of V, $\alpha\vec{x} \in W$, and $\alpha\vec{x} \in W_2$. Showing that $\alpha\vec{x} \in W_1 \cap W_2$. Hence $W_1 \cap W_2$ is closed under addition and scalar multiplication and thus $W_1 \cap W_2$ is a subspace of V.

$W_1 \cup W_2$ need not be a subspace of V. For this, consider $V = \mathbb{R}^3$ and let $W_1 = \{(a, b, 0) \mid a, b \in \mathbb{R}\}$ and $W_2 = \{(0, b, c) \mid b, c \in \mathbb{R}\}$. Then, clearly W_1 and W_2 are subspaces of $\mathbb{R}^3$. Now, $(2, 1, 0) \in W_1$ and $(0, 2, 4) \in W_2$, that is (2, 1, 0) and (0, 2, 4) belong to $W_1 \cup W_2$. But their sum $(2, 1, 0) + (0, 2, 4) = (2, 3, 4) \notin W_1 \cup W_2$, since (2, 3, 4) neither belongs W_1 nor belongs to W_2. So $W_1 \cup W_2$. is not closed under addition, hence cannot be a subspace of V.

Note : For any vector space V, $\{\vec{0}\}$ and V itself are subspaces of V. These subspaces are called *trivial subspaces* of V.

SOLVED EXAMPLES

Example 3.17 : Let $V = M_{2\times 2}(\mathbb{R})$ be the vector space of all 2×2 matrices with real entries; and W be the set of all matrices of the form $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ a, b $\in \mathbb{R}$. Then W is a subspace of V.

Solution : If $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and $B = \begin{bmatrix} a' & 0 \\ 0 & b' \end{bmatrix}$ are any two members in W, then,

$$A + B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} a' & 0 \\ 0 & b' \end{bmatrix}$$

$$= \begin{bmatrix} a + a' & 0 \\ 0 & b + b' \end{bmatrix} \in W$$

and for any scalar k, $kA = k\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

$$= \begin{bmatrix} ka & 0 \\ 0 & kb \end{bmatrix} \in W$$

which shows that both A + B and kA are again in W.

Therefore, W is a subspace of V.

Example 3.18 : Let V be a vector space of all real valued functions defined on closed interval [0, 1]. Let W be the set of all functions f in V for which f (c) > 0, where c is fixed number in [0, 1]. Is W a subspace of V? Justify your answer.

Solution : W is not a subspace of V, because if f is in W, then f (c) > 0 by given hypothesis. If we take k = –1, then [(–1) f] (c) = (–1) f (c) < 0 as f (c) > 0. Therefore, [(–1) f] (c) < 0, hence (–1) f cannot belong to W. So W is not closed under scalar multiplication, *hence is not a subspace*.

Example 3.19 : Let V be a vector space given in Example (3.18) above. Let

$$W = \left\{ f \in V \mid f\left(\tfrac{1}{2}\right) = 0 \right\}$$

Then show that W is a subspace of V.

Solution : For any f and g in W, we have, $f\left(\tfrac{1}{2}\right) = 0 = g\left(\tfrac{1}{2}\right)$, so that,

$$(f + g)\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right) + g\left(\frac{1}{2}\right), \text{ by sum in V.}$$
$$= 0 + 0$$

Thus, $(f + g)\left(\frac{1}{2}\right) = 0$, Hence f + g is in W.

Also for any scalar k, we have,

$$(kf)\left(\frac{1}{2}\right) = kf\left(\frac{1}{2}\right) \quad \text{by definition of scalar multiplication.}$$
$$= k \cdot 0, \quad \because f\left(\tfrac{1}{2}\right) = 0, \text{ as } f \in W$$
$$= 0 \quad \text{by theorem(2) (ii),}$$

Thus, $kf\left(\frac{1}{2}\right) = 0 \Rightarrow kf \in W$.

Thus, W is closed under addition and scalar multiplication, hence W is a subspace of V.

Example 3.20 : Let $V = \mathbb{R}^3$, the vector space of ordered triplets of real numbers. If $W = \{ (x, y, z) \in \mathbb{R}^3 \mid 2x + 3y + 4z = 0 \}$, show that W is a subspace of V.

Solution : Let $\vec{u} = (x, y, z)$, $\vec{v} = (x', y', z')$ be any vectors in W. Then by hypothesis for W, we have,

$$2x + 3y + 4z = 0 \quad \ldots(i)$$

and $$2x' + 3y' + 4z' = 0 \quad \ldots(ii)$$

Now, $\vec{u} + \vec{v} = (x + x', y + y', z + z')$, by definition of addition in V.

In order to show that u + v is in W, we need to show that,

$2 (x + x') + 3(y + y') + 4 (z + z') = 0$

But $2 (x + x') + 3 (y + y') + 4 (z + z')$

$= (2x + 3y + 4y) + (2x' + 3y' + 4z')$, by properties of real numbers.

$= 0 + 0$ by (i) and (ii)

$= 0$

$\therefore \vec{u} + \vec{v} \in W.$

Also for any scalar k to show ku is in W, that is, (kx, ky, kz) is in W, we have to show that,

$$2(kx) + 3(ky) + 4(kz) = 0$$

But $2(kx) + 3(ky) + 4(kz) = k(2x + 3y + 4z)$

$= k \cdot 0,$ by (i)

$= 0$

Thus, $k\vec{u}$ is in W. Hence, W is a subspace of V.

Example 3.21 : Let V be the function space as given in Example 3.9 of section 3.1. Let W be the set of all real polynomials in x, for x in [a, b]. Then W is a subspace of V.

Solution : We have already noted that W is a subset of V and W is itself a vector space, hence, is a subspace.

However, let us show that W is a subspace of V, by showing that W is closed under addition and scalar multiplication. For this, let

$$p(x) = a_0 + a_1x + \ldots\ldots\ldots + a_nx^n$$

and $$q(x) = b_0 + b_1x + \ldots\ldots\ldots + b_mx^m$$

be any two polynomials in W. Then, we have,

$$p(x) + q(x) = (a_0 + a_1x + \ldots\ldots\ldots + a_nx^n) + (b_0 + b_1x + \ldots\ldots\ldots + b_mx^m)$$

$$= (a_0 + b_0) + (a_1 + b_1)x + \ldots\ldots$$

which shows that p (x) + q (x) is in W. Also for any scalar k, we have,

$$kp(x) = k(a_0 + a_1x + \ldots\ldots\ldots a_nx^n)$$

$$= (ka_0) + (ka_1)x + \ldots\ldots\ldots + (ka_n)x^n.$$

Thus, kp (x) is also in W, showing that W satisfies the conditions of theorem 3, hence, is a subspace of V.

Example 3.22 : Let V be a function space as given in Example 3.9 and W be the set of all constant functions defined on [a, b]. Then show that W is a subspace of V.

Solution : Let f and g be in W, then by hypothesis $f(x) = c_1$ and $g(x) = c_2$ for all x in [a, b], where c_1 and c_2 are any fixed real numbers. Then

$$\forall x \in [a, b], (f + g)(x) = f(x) + g(x), \quad \text{by sum in V.}$$

$$= c_1 + c_2,$$

which shows that f + g is also constant function, hence belongs to W. Also, for any scalar k,

$$(kf)(x) = kf(x), \quad \text{for all x in [a, b]}$$

$$= k \cdot c_1, \quad \text{for all x in [a, b]}$$

thus kf is also constant function and hence, belongs to W, showing that W is a subspace of V.

Note : The subspace W in Example 3.22 above coincides with the whole set of real numbers.

Example 3.23 : Let $V = \mathbb{R}^3$ and W be the subset of V, given by,

$$W = \{ (x, y, z) \in V \mid x = 0 \text{ or } y = 0 \}.$$

Is W a subspace of V ? why ?

Solution : W is not subspace of V, because (0, 3, 5) and (5, 0, –1) are in W, whereas the sum (0, 3, 5) + (5, 0, –1) = (5, 3, 4) cannot be in W, since neither x is zero nor y is zero.

Example 3.24 : Let $V = \mathbb{R}^3$, and $W = \{(x, y, z) \in V | x^2 - y^2 = 0\}$. Is W a subspace of V ? Why ?

Solution : W is not a subspace of V, because W is not closed under addition. For $\vec{u} = (-3, 3, 1)$ and $\vec{v} = (4, +4, -1)$ are in W as $x^2 - y^2 = 0$ for both the vectors, but $\vec{u} + \vec{v} = (1, 7, 0)$ is not in W, as x = 1 and y = 7, showing that $x^2 - y^2 = 1 - 49 = -48 \neq 0$.

Example 3.25 : Show that the set W of all solutions of system of m-linear homogeneous equations in n-unknown forms a subspace of $\mathbb{R}^n$.

Solution : Any system of m-linear homogeneous equations in n-unknowns is given by :

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ldots + a_{2n}x_n = 0 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots\ldots + a_{mn}x_n = 0 \end{array}\right\} \quad \ldots(I)$$

This system can be written precisely in matrix form as,

$$AX = 0 \quad \ldots(II)$$

where, $A = (a_{ij})_{m \times n}$ is matrix of coefficients and X in $n \times 1$ column matrix of unknowns. Any solution to (I) or (II) is an $n \times 1$ column matrix of real numbers, which can be considered as a vector in $\mathbb{R}^n$. That is, the set W of solutions of system (I) is a subset of $\mathbb{R}^n$. The trivial solution belongs to W, and hence, W is non-empty.

Now, to show that W is subspace of $\mathbb{R}^n$ we have to prove that W is closed under addition and scalar multiplication. Let X_1 and X_2 be any two solutions of the system (II), then

$$AX_1 = 0 \text{ and } AX_2 = 0.$$

so that,
$$\begin{aligned} A(X_1 + X_2) &= AX_1 + AX_2 \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

Showing that $X_1 + X_2$ is also a solution of (II). Also $A(kX_1) = k(AX_1) = k \cdot 0 = 0$.

showing that kX_1 is in W. Thus, W is subspace of $\mathbb{R}^n$.

This subspace W is called the *solution space of system of homogeneous linear equations*.

Example 3.26 : From examples 10, 11, 12, 13 and 14 in 3.1, we see that

(i) C ([a, b]) is a subspace of F ([a, b], $\mathbb{R}$).

(ii) D ([a, b]) is a subspace of F ([a, b], $\mathbb{R}$].

Also, we know that differentiable functions are continuous, D ([a, b]) is subset of C ([a, b]), hence D ([a, b]) is a subspace of C ([a, b]).

(iii) $\mathbb{R}$ ([a, b]) is subspace of F ([a, b], $\mathbb{R}$).

(iv) P is a subspace of F ([a, b), $\mathbb{R}$] if $x \in [a, b]$.

(v) P_n is a subspace of P.

Example 3.27 : The set S_n of symmetric real matrices and the set A_n of antisymmetric matrices are subspace of M (n, $\mathbb{R}$).

Example 3.28 : Let V be a vector space and let $\vec{v} \in V$. Then, the smallest subspace of V containing $\vec{v}$ is

$$V_0 = \{\alpha \vec{v} \mid \alpha \in \mathbb{R}\}.$$

Solution : If $\vec{x} = \alpha \vec{v} \in V_0$, $\vec{y} = \beta \vec{v} \in V_0$, for $\alpha, \beta \in \mathbb{R}$, then $\vec{x} + \vec{y} = \alpha \vec{v} + \beta \vec{v}$ $= (\alpha + \beta)\ \vec{v} = \gamma \vec{v}$, $\in V_0$, where $\gamma = \alpha + \beta$. Similarly, if $\vec{u} \in V_0$ is any vector in V_0, then for any scalar δ, $\delta \vec{u} \in V_0$, because $\vec{u} = \alpha \vec{v}$, so $\delta \vec{u} = (\delta\alpha)\ \vec{v} = \gamma \vec{v}$, where $\gamma = \delta\alpha$. Thus, V_0 is closed under addition and scalar multiplication, hence is a subspace of V. Since $\vec{v} = 1 \cdot \vec{v} \in V_0$, hence $\vec{v}$ is contained in V_0. If W is any subspace of V containing $\vec{v}$, by definition of subspace, every scalar multiple $\alpha \vec{v} \in W$, hence $V_0 \subseteq W$. Therefore, V_0 is the smallest subspace of V containing $\vec{v}$.

Example 3.29 : Let $\vec{v} = (r, s) \in \mathbb{R}^2$ be a non-zero vector in $\mathbb{R}^2$. Then, by example (3.28), the smallest subspace V_0 of $\mathbb{R}^2$ containing $\vec{v}$ is the set

$$V_0 = \{t \vec{v} \mid t \in \mathbb{R}\} = \{(tr, ts) \mid t \in \mathbb{R}\}.$$

This is nothing but the line through the origin and the point (r, s) in $\mathbb{R}^2$. As we know the line joining the origin and (r, s) is given by the equation

$$\frac{x - 0}{0 - r} = \frac{y - 0}{0 - s} = -t \text{ say}$$

Hence, any point on this line is given by (tr, ts) for some $t \in \mathbb{R}$.

Note : Thus, we see from example (3.29) that any line passing through the origin is a subspace of $\mathbb{R}^2$. Infact, converse is also true; that is, any non-trivial subspace of $\mathbb{R}^2$ is a line passing through the origin.

Example 3.30 : If a non-zero vector $\vec{v} = (r, s, t) \in \mathbb{R}^3$, then the smallest subspace V_0 of $\mathbb{R}^3$ containing $\vec{v}$ is the set $V_0 = \{\alpha (r, s, t) \mid \alpha \in \mathbb{R}\} = \{(\alpha r, \alpha s, \alpha t) \mid \alpha \in \mathbb{R}\}$.
which is clearly a line passing through origin and the point (r, s, t).

Solution : Similar to example 3.28.

Note : The lines passing through the origin are subspaces of $\mathbb{R}^3$.

Example 3.31 : Let $\vec{v}$, $\vec{w}$ be any two vectors in $\mathbb{R}^3$. Let V_0 be subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$. As V_0 is subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$, by definition of subspace, $\alpha\vec{v} \in V_0$, and $\beta\vec{w} \in V_0$ for any $\alpha, \beta \in \mathbb{R}$. Again as V_0 is closed under addition, $\alpha\vec{v} + \beta\vec{w} \in V_0$. We claim that

$$V_0 = \{\alpha\vec{v} + \beta\vec{w} \mid \alpha, \beta \in \mathbb{R}\}$$

is the smallest subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$.

Since, for $\alpha = 1, \beta = 0$, $1\cdot\vec{v} + 0\cdot\vec{w} = \vec{v} \in V_0$ and $w = 0\cdot\vec{v} + 1\cdot\vec{w} \in V_0$, we see that V_0 contains $\vec{v}$ and $\vec{w}$.

To show that V_0 is subspace of $\mathbb{R}^3$:

Let $\vec{x}$, $\vec{y}$ be any vector in V_0, then

$\vec{x} = \alpha_1\vec{v} + \beta_1\vec{w}$ and $\vec{y} = \alpha_2\vec{v} + \beta_2\vec{w}$, for some scalars $\alpha_1, \beta_1, \alpha_2, \beta_2$ in $\mathbb{R}$. Then,

$$\begin{aligned}
\vec{x} + \vec{y} &= (\alpha_1\vec{v} + \beta_1\vec{w}) + (\alpha_2\vec{v} + \beta_2\vec{w}) \\
&= (\alpha_1\vec{v} + \alpha_2\vec{v}) + (\beta_1\vec{w} + \beta_2\vec{w}) && \text{by vector space axioms.} \\
&= (\alpha_1 + \alpha_2)\vec{v} + (\beta_1 + \beta_2)\vec{w} && \text{by vector space axioms.} \\
&= \alpha_3\vec{v} + \beta_3\vec{w} \in V_0
\end{aligned}$$

where $\alpha_3 = \alpha_1 + \alpha_2, \beta_3 = \beta_1 + \beta_2$

Also for any scalar α,

$$\begin{aligned}
\alpha\vec{x} &= \alpha(\alpha_1\vec{v} + \beta_1\vec{w}) = \alpha(\alpha_1\vec{v}) + \alpha(\beta_1\vec{w}) \\
&= (\alpha\alpha_1)\vec{v} + (\alpha\beta_1)\vec{w} \\
&= \alpha_4\vec{v} + \beta_4\vec{w} \in V_0
\end{aligned}$$

where, $\alpha_4 = \alpha\alpha_1$ and $\beta_4 = \alpha\beta_1$.

Hence, V_0 is a subspace of $\mathbb{R}^3$; containing $\vec{v}$ and $\vec{w}$.

To show that V_0 is the smallest subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$. Suppose W is a subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$. Since W is subspace containing $\vec{v}$ and $\vec{w}$, $\alpha\vec{v} \in W$ and $\beta\vec{w} \in W$, for any scalars α, β, hence $\alpha\vec{v} + \beta\vec{w} \in W$ for any α, β in $\mathbb{R}$.

Hence, $V_0 \subseteq W$.

Example 3.32 : Let $\vec{v} = (r, s, t) \in \mathbb{R}^3$ and $\vec{w} = (a, b, c) \in \mathbb{R}^3$ be two non-zero vectors. Then, from above example (3.31), the smallest subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$ is the set,

$$V_0 = \{\alpha\vec{v} + \beta\vec{w} \mid \alpha, \beta \in \mathbb{R}\}.$$

This is nothing but a plane containing the origin, $\vec{v}$ and $\vec{w}$, provided that $\vec{v}$ and $\vec{w}$ are not scalar multiples of each other. To see this we know that the equation of a plane through origin (0, 0, 0) and points (r, s, t), (a, b, c) is given by,

$$\begin{vmatrix} x & y & z & 1 \\ r & s & t & 1 \\ a & b & c & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 0$$

or

$$\begin{vmatrix} x & y & z \\ r & s & t \\ a & b & c \end{vmatrix} = 0$$

or $(bt - sc)\, x + (at - rc)\, y + (as - rb)\, z = 0$

or $a'x + b'y + c'z = 0$, the equation of the plane through origin and containing $\vec{v}$ and $\vec{w}$.

Thus V_0 is a plane in $\mathbb{R}^3$ containing $\vec{v}$, $\vec{w}$ and origin, if $\vec{v}$ and $\vec{w}$ are not scalar multiples of each other.

If $\vec{v}$ and $\vec{w}$ are scalar multiples of each other, then either $\vec{v} = \alpha\vec{w}$ or $\vec{w} = \beta\vec{v}$ for some α, β in $\mathbb{R}$. In either case V_0 becomes,

$V_0 = \{\alpha v \mid \alpha \in \mathbb{R}\}$ which is a line through origin and containing $\vec{v}$.

Thus, we conclude : If $\vec{v} = (r, s, t)$, $\vec{w} = (a, b, c)$ are two non-zero vectors in $\mathbb{R}^3$, then the smallest subspace of $\mathbb{R}^3$ containing $\vec{v}$ and $\vec{w}$ is either

(i) a plane containing these points and the origin or

(ii) a line passing through these points and the origin, where $\vec{v}$ and $\vec{w}$ are scalar multiples of each other.

Remark :

From examples (3.29) and (3.32), we remark :

(i) The only subspaces of $\mathbb{R}^2$ are $\{0\}$ and all the lines passing through the origin.

(ii) The only subspaces of $\mathbb{R}^3$ are $\{0\}$, all the lines passing through origin and all the planes passing through origin.

3.3 Linear Combination and Spanning Sets

Definition : Let V be a vector space and $\vec{v_1}, \vec{v_2}, \ldots\ldots \vec{v_r}$ be vectors in V. An element $\vec{u}$ in V of the form $\vec{u} = k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots\ldots + k_r\vec{v_r}$, where $k_1, k_2, \ldots\ldots\ldots, k_r$ are scalars, is called a linear combination of the vectors $\vec{v_1}, \vec{v_2}, \ldots\ldots\ldots, \vec{v_r}$.

SOLVED EXAMPLES

Example 3.33 : Let $\vec{v_1} = (1, 2, 3)$ and $\vec{v_2} = (0, -1, 2)$ be vectors in $\mathbb{R}^3$. Show that : (a) $(-1, -4, 1)$ is a linear combination of $\vec{v_1}$ and $\vec{v_2}$. (b) $(4, 7, 15)$ is not linear combination of $\vec{v_1}$ and $\vec{v_2}$.

Solution : (a) In order to show that $\vec{u} = (-1, -4, 1)$ is linear combination of $\vec{v_1}$ and $\vec{v_2}$, there must be scalars k_1 and k_2 such that,

$$(-1, -4, 1) = k_1\vec{v_1} + k_2\vec{v_2}$$

or $$(-1, -4, 1) = k_1 (1, 2, 3) + k_2 (0, -1, 2)$$

or $$(-1, -4, 1) = (k_1, 2k_1 - k_2, 3k_1 + 2k_2)$$

or on equating the corresponding components, we have,

$$\begin{aligned} k_1 &= -1. \\ 2k_1 - k_2 &= -4 \\ 3k_1 + 2k_2 &= 1 \end{aligned}$$

Solving this system we get $k_1 = -1$ and $k_2 = 2$, which satisfy all the above three equations. Thus, we have,

$$\vec{u} = (-1, -4, 1) = (-1)\,\vec{v_1} + 2 \cdot \vec{v_2}.$$

Hence, u is linear combination of $\vec{v_1}$ and $\vec{v_2}$.

(b) Similarly for $\vec{u} = (4, 7, 15)$, suppose there are scalars k_1 and k_2 such that,

$$(4, 7, 15) = k_1(1, 2, 3) + k_2 (0, -1, 2) \qquad \ldots(i)$$

or $$(4, 7, 15) = (k_1, -2k_1 - k_2, 3k_1 + 2k_2)$$

Equating the corresponding components we get,

$$\begin{aligned} k_1 &= 4 \\ 2k_1 - k_2 &= 7 \qquad (*) \\ 3k_1 + 2k_2 &= 15. \end{aligned}$$

Solving the first two equations we get, $k_1 = 4$ and $k_2 = 1$, but these values of k_1 and k_2 do not satisfy the third equation of system (*), hence it is inconsistent. That is, there are

no such scalars k_1 and k_2, satisfying the identity (i). Consequently, $\vec{u}$ is not linear combination of $\vec{v_1}$ and $\vec{v_2}$.

Example 3.34 : (a) Let $\vec{v_1} = (1, 0)$ and $\vec{v_2} = (0, 1)$ be two vectors in $\mathbb{R}^2$. Show that every vector in $\mathbb{R}^2$ is a linear combination of $\vec{v_1}$ and $\vec{v_2}$.

(b) Let $\vec{u} = (1, 1, 1)$, $\vec{v} = (1, 1, 0)$ and $\vec{w} = (1, 0, 0)$ be three vectors in $\mathbb{R}^3$. Show that (3, 2, 1) is linear combination of $\vec{u}$, $\vec{v}$ and $\vec{w}$.

Solution : (a) Let (a, b) be arbitrary vector in $\mathbb{R}^2$.

Then, we can express (a, b) as :

$$(a, b) = a(1, 0) + b(0, 1)$$

which shows that (a, b) is a linear combination of (1, 0) and (0, 1). Hence, every vector in $\mathbb{R}^2$ is a linear combination of (1, 0) and (0, 1).

(b) To show that (3, 2, 1) is linear combination of (1, 1, 1), (1, 1, 0) and (1, 0, 0), we need to show that there are scalars k_1, k_2 and k_3 such that,

$$(3, 2, 1) = k_1(1, 1, 1) + k_2(1, 1, 0) + k_3(1, 0, 0)$$

or $(3, 2, 1) = (k_1 + k_2 + k_3, k_1 + k_2, k_1)$, by scalar multiplication and addition in $\mathbb{R}^3$.

Equating the corresponding components, we get,

$$\begin{aligned} k_1 + k_2 + k_3 &= 3 \\ k_1 + k_2 &= 2 \\ k_1 &= 1 \end{aligned}$$

Solving these equations for k_1, k_2 and k_3 we obtain $k_1 = 1$, $k_2 = 1$ and $k_3 = 1$, so that,

$$(3, 2, 1) = 1(1, 1, 1) + 1(1, 1, 0) + 1(1, 0, 0).$$

Note : In case of (b) above, every vector in $\mathbb{R}^3$ is also linear combination of (1, 1, 1), (0, 1, 1) and (0, 0, 1). (Work out as an exercise).

Definition : Let V be a vector space and $S = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}\}$ be the set of vectors in V. Then the set of all linear combinations of the vectors $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$ is called a linear span of the set S and is denoted by L (S) or lin (S). Thus,

$$L(S) = \left\{k_1\vec{v_1} + k_1\vec{v_2} + \ldots\ldots + k_r\vec{v_r} \mid k_1, k_2, \ldots\ldots, k_r \text{ are scalars}\right\}$$

Note :

(1) We also say sometimes that L (S) is spanned by the set S.

(2) From above definition, we note that $\mathbb{R}^2$ is a linear span of just two vectors (1, 0) and (0, 1). Similarly one can observe that $\mathbb{R}^3$ is a linear span of the vectors (1, 1, 1), (1, 1, 0) and (1, 0, 0).

(3) Linear span of the empty set is defined to be zero vector, that is, $L(\phi) = \{0\}$.

Theorem 4 : Let $S = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}\}$ be a set of vectors in a vector space V. Then

(a) L (S) is a subspace of V.

(b) L (S) is the smallest subspace of V containing S.

Proof : (a) In order to show that L (S) is a subspace of V, we must prove that L (S) is closed under addition and scalar multiplication. If $\vec{u}$ and $\vec{v}$ are vectors in L (S), then,

$$\vec{u} = k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots\ldots + k_r\vec{v_r}$$

and $$\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + \ldots\ldots\ldots + c_r\vec{v_r},$$

where, $k_1, k_2, \ldots\ldots, k_r$ and $c_1, c_2, \ldots\ldots\ldots, c_r$ are scalars.

$$\therefore \quad \vec{u} + \vec{v} = (k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots + k_r\vec{v_r}) + (c_1\vec{v_1} + c_2\vec{v_2} + \ldots\ldots + c_r\vec{v_r})$$

$$= (k_1 + c_1)\vec{v_1} + (k_2 + c_2)\vec{v_2} + \ldots\ldots + (k_r + c_r)\vec{v_r}$$

using vector space axioms.

Also, for any scalar k,

$$k\vec{u} = k(k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots + k_r\vec{v_r})$$

$$= (kk_1)\vec{v_1} + (kk_2)\vec{v_2} + \ldots\ldots\ldots + (kk_r)\vec{v_r}.$$

Thus, $\vec{u} + \vec{v}$ and $k\vec{u}$ are linear combinations of $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$ and consequently are members of L(S). Thus, L(S) is closed under addition and scalar multiplication. Therefore L (S) is a subspace.

(v) First note that each vector $\vec{v_i}$ is a linear combination of $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$, since we can write $\vec{v_i} = 0.\vec{v_1} + 0.\vec{v_2} + \ldots\ldots + 1\vec{v_i} + \ldots\ldots + 0\vec{v_r}$. Therefore, L(S) contains each of the vectors $\vec{v_1}, \vec{v_2}, \ldots\ldots \vec{v_r}$. If W is any other subspace of V containing $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$, it must contain all linear combinations of $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$ as it is subspace so that W contains each of the vectors in L(S), hence, $L(S) \subseteq W$. Thus, any subspace of V containing $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$ also contains L(S). Therefore, L (S) is the smallest subspace of V containing S.

SOLVED EXAMPLES

Example 3.35 : Show that the vectors

$$\vec{v_1} = (1, 1, 1), \vec{v_2} = (0, 1, 1) \text{ and } \vec{v_3} = (0, 1, -1) \text{ span } \mathbb{R}^3.$$

Solution : In order to show that $\overrightarrow{v_1}$, $\overrightarrow{v_2}$, and $\overrightarrow{v_3}$ span $\mathbb{R}^3$, we have to prove that each vector in $\mathbb{R}^3$ is a linear combination of $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ and $\overrightarrow{v_3}$. If (a, b, c) is any vector in $\mathbb{R}^3$, we have to prove that there exist scalars k_1, k_2 and k_3 such that

$$(a, b, c) = k_1 (1, 1, 1) + k_2 (0, 1, 1) + k_3 (0, 1, -1).$$

This gives using addition and scalar multiplication :

$$(a, b, c) = (k_1, k_1 + k_2 + k_3, k_1 + k_2 - k_3).$$

Equating the corresponding components, we get,

$$\begin{aligned} k_1 &= a \\ k_1 + k_2 + k_3 &= b \\ k_1 + k_2 - k_3 &= c \end{aligned}$$

Solving these equations for k_1, k_2 and k_3, we obtain $k_1 = a$, $k_2 = \frac{1}{2}(b + c - 2a)$, $k_3 = \frac{1}{2}(b - c)$. Thus, (a, b, c) is a linear combination of $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ and $\overrightarrow{v_3}$. Since (a, b, c) is arbitrary vector in $\mathbb{R}^3$, $\mathbb{R}^3 \subseteq L(S)$. But $L(S) \subseteq \mathbb{R}^3$, thus we have proved that $L(S) = \mathbb{R}^3$.

Example 3.36 : Show that the vectors $\overrightarrow{v_1} = (2, -1, 3)$, $\overrightarrow{v_2} = (4, 1, 2)$ and $\overrightarrow{v_3} = (8, -1, 8)$ do not span $\mathbb{R}^3$.

Solution : If $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ and $\overrightarrow{v_3}$ span $\mathbb{R}^3$, for any vector (a, b, c) in $\mathbb{R}^3$ there must exist scalars k_1, k_2 and k_3 such that,

$$(a, b, c) = k_1 \overrightarrow{v_1} + k_2 \overrightarrow{v_2} + k_3 \overrightarrow{v_3} \quad \ldots(*)$$

or $$(a, b, c) = (2k_1 + 4k_2 + 8k_3, -k_1 + k_2 - k_3, 3k_1 + 2k_2 + 8k_3).$$

Equating corresponding components, we get,

$$\begin{aligned} 2k_1 + 4k_2 + 8k_3 &= a \quad \ldots(i) \\ -k_1 + k_2 - k_3 &= b \quad \ldots(ii) \\ 3k_1 + 2k_2 + 8k_3 &= c \quad \ldots(iii) \end{aligned}$$

Solving by elimination,

Multiply equation (i) by $\frac{1}{2}$ and then add it to equation (ii) and substract resulting equation three times from equation (iii), we obtain,

$$\begin{aligned} k_1 + 2k_2 + 4k_3 &= \frac{a}{2} \\ 3k_2 + 3k_3 &= \left(\frac{a}{2} + 3\right) \\ -4k_2 - 4k_3 &= \left(c - \frac{3}{2}a\right) \end{aligned}$$

or $$k_1 + 2k_2 + 4k_3 = \frac{a}{2} \quad \ldots(i)$$

$$k_2 + k_3 = \frac{1}{3}\left(\frac{a}{2} + b\right) \quad \ldots(ii)$$

$$k_2 + k_3 = -\frac{1}{4}\left(c - \frac{3}{2}a\right) \quad \ldots(iii)$$

The left hand sides of equations (ii) and (iii) are equal, whereas the right hand sides of the same equations need not be always equal depending upon the values of a, b and c. This shows that no scalars satisfying (∗) can exist. This is same thing as that not every vector in $\mathbb{R}^3$ is a linear combination of $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$. Hence, $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$ cannot span $\mathbb{R}^3$.

Example 3.37 : Show that the polynomials $P_1 = 1 - x + 2x^2$; $P_2 = 3 + x$; $P_3 = 5 - x + 4x^2$; $P_4 = -2 - 2x + 2x^2$ do not span the vector space P_2 of polynomials of degree ≤ 2.

Solution : Any polynomial of degree ≤ 2 can be written as $f(x) = a_0 + a_1x + a_2x^2$, where a_0, a_1, and a_2 are scalars. If the polynomials P_1, P_2, P_3 and P_4 span P_2, there must exist scalars k_1, k_2, k_3 and k_4 such that,

$$f(x) = k_1P_1 + k_2P_2 + k_3P_3 + k_4P_4 \quad \ldots(I)$$

or $a_0 + a_1x + a_2x^2 = k_1(1 - x + 2x^2) + k_2(3 + x) + k_3(5 - x + 4x^2) + k_4(-2 - 2x + 2x^2)$.

or $a_0 + a_1x + a_2x^2 = (k_1 + 3k_2 + 5k_3 - 2k_4) + (-k_1 + k_2 - k_3 - 2k_4)x + (2k_1 - 4k_3 + 2k_4)x^2$

Equating the coefficients on both sides we get,

$$\left.\begin{aligned} k_1 + 3k_2 + 5k_3 - 2k_4 &= a_0 \\ -k_1 + k_2 - k_3 - 2k_4 &= a_1 \\ 2k_1 + 4k_3 + 2k_4 &= a_2 \end{aligned}\right\} \quad \ldots(*)$$

If we try to solve this system (∗) by Gauss-elimination method, we obtain a condition $2a_2 + 3a_1 - a_0 = 0$ on a_0, a_1, and a_2, which will not be satisfied for all choices of a_0, a_1 and a_2, showing that the system (∗) is inconsistent, that is, the system (∗) has no solution for all choices of $a_0 + a_1x + a_2x^2$. Thus, there does not exist scalars k_1, k_2, k_3 and k_4 satisfying (∗) simultaneously, hence, the identity (I) showing that P_1, P_2, P_3 and P_4 cannot span the vector space P_2.

Example 3.38 : Let S and T be subsets of a vector space V. Then,

(a) show that if $S \subset T$, then $L(S) \subset L(T)$.

(b) show that $L(S) = S$, if and only if, S is a subspace of V.

(c) show that $L(L(S)) = L(S)$.

Solution : (a) If $\vec{u} \in L(S)$ is any vector, then $\vec{u} = k_1\vec{v_1} + k_2\vec{v_2} + \ldots + k_r\vec{v_r}$, where, $k_1, k_2, \ldots, k_r$ are scalars and $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_r}$ are vectors in S. As $S \subset T$; $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_r}$ also

belong to T and hence their linear combination $\vec{u}$ belongs to L (T), that is, $\vec{u} \in L(T)$. Therefore, $L(S) \subset L(T)$.

(b) If L(S) = S, then we know that L(S) is a subspace of V, hence, S is subspace of V.

Conversely, if S is a subspace of V, we have to show that L(S) = S. For this, if $\vec{u} \in L(S)$, then $\vec{u} = k_1 \vec{v_1} + \ldots + k_r \vec{v_r}$, where $k_1, k_2, \ldots, k_r$ are scalars and $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_r}$ are in S. But S is a subspace, so that it is closed under addition and scalar multiplication, hence, $\vec{u}$ must be in S, which shows that

$$L(S) \subseteq S \qquad \ldots(i)$$

On the other hand, if $\vec{v} \in S$ be any vector, $\vec{v}$ can be written as linear combination of vectors in S as :

$$\vec{v} = 1 \cdot \vec{v} + 0.\vec{v_1} + 0.\vec{v_2} + \ldots\ldots + 0.\vec{v_r},$$

where, $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r} \in S$, which shows that $\vec{v}$ is linear combination of vectors in S. Hence, $\vec{v} \in L(S)$, so that

$$S \subseteq L(S) \qquad \ldots(ii)$$

From (i) and (ii) we get,

$$L(S) = S$$

(c) By using the fact in (b), and we know that L (S) is a subspace of V, we get,

$$L(L(S)) = L(S)$$

Definition : A vector space V is said to be **finitely generated** if there is a finite set S such that V = L(S).

Result 1 : Let $S = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be the set of vectors in a vector space V and let W = L(S) − linear span of S. If some $\vec{u_i}$, $1 \le i \le n$ is linear combination of the remaining vectors of S, then the set $S_1 = \{\vec{u_1}, \ldots, \vec{u_{i-1}}, \vec{u_{i+1}}, \ldots, \vec{u_n}\}$ also spans W; that is $W = L(S_1)$.

Proof : Since $\vec{u_i}$ is linear combination of rest of the vectors in S, then we have

$$\vec{u_i} = \alpha_1 \vec{u_1} + \ldots + \alpha_{i-1} \vec{u_{i-1}} + \alpha_{i+1} \vec{u_{i+1}} + \ldots + \alpha_n \vec{u_n} \qquad \ldots (i)$$

(for some scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$)

If $\vec{w}$ is any vector in W and as W = L(S), we have, scalars $a_1, a_2, \ldots, a_n$ such that

$$\vec{w} = a_1\vec{u}_1 + \ldots + a_{i-1}\vec{u}_i + a_i\vec{u}_i + a_{i+1}\vec{u}_{i+1} + \ldots + a_n\vec{u}_n$$

Replacing $\vec{u}_i$ by expression of $\vec{u}_i$ in (1), we have

$$\vec{w} = a_1\vec{u}_1 + \ldots + a_{i-1}\vec{u}_{i-1} + a_i(\alpha_1\vec{u}_1 + \ldots \alpha_{i-1}\vec{u}_{i-1} + \alpha_{i+1}\vec{u}_{i+1} + \ldots + \alpha_n\vec{u}_3)$$

$$+ a_{i+1}\vec{u}_{i+1} + \ldots + a_n\vec{u}_n$$

Using vector space axioms rigously we get

$$\vec{w} = (a_1 + a_i\alpha_1)\vec{u}_1 + \ldots + (a_{i-1} + a_i\alpha_{i-1})\vec{u}_{i-1}$$

$$+ (a_i\alpha_{i+1} + a_{i+1})\vec{u}_{i+1} + \ldots + (a_i\alpha_n + a_n)\vec{u}_n$$

Showing that $\vec{w}$ is a linear combination of the vectors in S_1, as $\vec{w}$ is arbitrary vector in W, we have $W = L(S_1)$.

Remark : The above result states that if $W = L(S)$ and one of the vector in S is linear combination of the vectors in S, then we can drop this vector from spanning set S and still retain a set which spans W.

Note : If S is a set of vectors in a vector space V and L(S) is the linear span of the set S. Here it is to be noted that S is not unique, in the sense that L(S) can be spanned by any other set not equal to S. For example, as we have seen that the lines passing through origin are subspaces of $\mathbb{R}^2$, this line can be spanned by any non-zero vector on it. Similarly, a plane in $\mathbb{R}^3$ passing through origin can be spanned by any two lines in it. The following is the theorem, which is the generalization of the above facts.

Theorem 5 : If $S = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$ and $S' = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ are two sets of vectors in a vector space V, then $L(S) = L(S')$ if and only if each of the vectors in S is a linear combination of those in S' and each of the vectors in S' is a linear combination of those in S.

Proof : If $L(S) = L(S')$, then clearly vector in S is linear combination of those in S' and any vector in S' is a linear combination of those in S. Conversely, let us suppose that each vector in S is a linear combination of those in S' and any vector in S' is a linear combination of those in S, we have to show that $L(S) = L(S')$.

For this suppose $\vec{u} \in L(S)$ be any vector, then by definition

$$\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots + \alpha_r\vec{u}_r \quad \ldots (i)$$

For some scalar $\alpha, \alpha_2, \ldots, \alpha_r$. By hypothesis each vector $\vec{u}_i$ in S is a linear combination of those in S', then we have

$$\vec{u}_i = \beta_{i1}\vec{v}_1 + \beta_{i2}\vec{v}_2 + \ldots + \beta_{ik}\vec{v}_k, \quad i = 1, \ldots r, \quad \ldots (ii)$$

For some scalars $\beta_{i1}, \beta_{i2}, \ldots, \beta_{ik}$. Using expression (ii) for each $\vec{u}_i$, $i = 1, 2, \ldots, r$ in (i) and using vector space axioms, we see that $\vec{u}$ is a linear combination of vectors in S', hence $\vec{u} \in L(S')$. But $\vec{u}$ is arbitrary vector in L(S), therefore $L(S) \subseteq L(S')$.

Similarly, reversing the arguments we can show that $L(S') \subseteq L(S)$, so $L(S) = L(S')$.

Remark : The above theorem is a generalization of the result 1.

Exercise 3.2

1. Let $V = \mathbb{R}^3$ be the vector space of ordered triples of real numbers with usual addition and scalar multiplication. Determine which of the following subjsets are subspaces of V. For those that are not, give reasons.
   (a) $W = \{(x, y, z) \mid x = 1, y = 1\}$
   (b) $W = \{(x, y, z) \mid x + y + z = 0\}$
   (c) $W = \{(x, y, z) \mid x + y + z = 1\}$
   (d) $W = \{(x, y, z) \mid x^2 - y^2 = 0\}$
   (e) $W = \{(x, y, z) \mid y = 3x, z = 5x\}$
   (f) $W = \{(x, y, z) \mid x + y = 0 \text{ and } x + 2y - 3z = 0\}$
   (g) $W = \{(x, y, z) \mid x \geq 0\}$
   (h) $W = \{(x, y, z) \mid \text{either } x = y \text{ or } y = z\}$
2. Let V be a function space of real valued functions defined on $[-1, 1]$ under pointwise addition and scalar multiplication. Determine which of the following subsets of V are subspaces of V. For those which are not, give reasons.
   (a) $W = \{f \in V \mid f(0) = 0\}$
   (b) $W = \{f \in V \mid f \text{ is an even function}\}$
   (c) $W = \{f \in V \mid f \text{ is an odd function}\}$
   (d) $W = [f \in V \mid f\left(\frac{1}{2}\right) > 0\}$
   (e) $W = \{f \in V \mid f \text{ is a bounded function}\}$
   (f) $W = \{f \in V \mid f\left(-\frac{1}{2}\right) = 1 + f\left(\frac{1}{2}\right)\}$.
   (g) $W = \{f \in V \mid f \text{ is a polynomial function of degree} \leq 3.\}$.
   (h) $W = \{f \in V \mid f(x) \leq 0, x \in [-1, 1]\}$.
3. Let V be the vector space of 2×2 real matrices with matrix addition and scalar multiplication. Determine which of the following subsets of V are subspaces of V. For those which are not, give reasons.
   (a) $W = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in V \mid a, b, c, d \text{ are integers}\right\}$
   (b) $W = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in V \mid a, b, c, d \text{ are rational numbers}\right\}$

(c) $W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in V \mid a \text{ is rational number} \right\}$

(d) $W = \{ A \in V \mid A^t = A \}$.

(e) $W = \{ A \in V \mid A^t = -A \}$.

(f) $W = \{ A \in V \mid \det A = 0 \}$.

(g) $W = \{ A \in V \mid A \text{ is non-singular} \}$.

(h) $W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in V \mid c = d = 0 \right\}$.

4. Let W_1 and W_2 be subspaces of a vector space V. Then show that,

   (a) $W_1 \cap W_2$ is a subspace of V.

   (b) $W_1 \cup W_2$ is subspace of V if and only if, either $W_1 \subseteq W_1$ or $W_2 \subseteq W_1$

   (c) $W_1 + W_2 = \{\overrightarrow{w_1} + \overrightarrow{w_2} \mid \overrightarrow{w_1} \in W_1 \text{ and } \overrightarrow{w_2} \in W_2\}$ is a subspace of V.

5. Let V be a vector space and $\overrightarrow{u}$, $\overrightarrow{v}$ are in V. Then show that the set $W = \{a\overrightarrow{u} + b\overrightarrow{v} \mid$ a, b are scalars$\}$ is a subspace of V.

6. Let A be an $m \times n$ matrix and $AX = 0$ be the homogeneous system of linear equations in n-unknowns and $W = \{X \in \mathbb{R}^n \mid AX = 0\}$. Then show that W is a subspace of $\mathbb{R}^n$.

   (Note here X the ordered n-tupple of solutions is considered to be an $n \times 1$ column matrix.)

7. Which of the following vectors are linear combinations of $\overrightarrow{u} = (-1, 1, 0)$ and $\overrightarrow{v} = (2, 1, 3)$? Justify your answer.

   (a) (1, 2, 3) ; (b) (7, 2, 9) ; (c) (5, 6, 3) ; (d) (0, 0, 0).

8. Express the following as linear combination of $\overrightarrow{u} = (-2, 1, 3)$, $\overrightarrow{v} = (3, 1, -1)$ and $\overrightarrow{w} = (-1, -2, 1)$?

   (a) (–6, –2, 6) ; (b) (8, 6, –6) ; (c) (0, 0, 0) ; (d) (6, –7, 9).

9. Express the following as linear combination of $P_1 = 1 - x + 2x^2$ and $P_2 = 3 + 2x - x^2$.

   (a) $7 + 8x - 7x^2$; (b) $-8 - 7x + 5x^2$; (c) $-4 - x - x^2$.

10. Which of the following are linear combinations of

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ -2 & 4 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} ?$$

   (a) $\begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}$; (b) $\begin{bmatrix} 3 & 1 \\ -4 & 6 \end{bmatrix}$; (c) $\begin{bmatrix} -4 & -2 \\ 0 & -2 \end{bmatrix}$.

11. In each part determine whether the given vectors span $\mathbb{R}^3$.

    (a) $\overrightarrow{u_1} = (1, 1, 1)$; $\overrightarrow{u_2} = (0, 1, 1)$; $\overrightarrow{u_3} = (0, 1, 1)$

    (b) $\overrightarrow{u_1} = (1, 2, 3)$; $\overrightarrow{u_2} = (0, 0, 1)$; $\overrightarrow{u_3} = (0, 1, 2)$

(c) $\vec{u_1} = (2, 1, 0)$; $\vec{u_2} = (0, 3, -4)$; $\vec{u_3} = (1, -1, 2)$

(d) $\vec{u_1} = (6, 2, 1)$; $\vec{u_2} = (-1, 2, 4)$; $\vec{u_3} = (5, 4, 5)$; $\vec{u_4} = (4, 6, 9)$.

12. Write the vector $\vec{v} = (1, -2, 5)$ as a linear combination of the vectors $\vec{u_1} = (1, 1, 1)$; $\vec{u_2} = (1, 2, 3)$ and $\vec{u_3} = (2, -1, 1)$.

13. For which value of k will the vector $\vec{v} = (1, -2, k)$ in $\mathbb{R}^3$ be a linear combination of the vectors $\vec{u} = (3, 0, -2)$ and $\vec{w} = (2, -1, -5)$?

Hints and Answers (3.2)

1. (a) No, $(0, 0, 0) \notin$ W; (b) Yes, (c) No, $(0, 0, 0) \notin$ W.
   (d) No, not closed under addition (e) Yes. (f) Yes.
   (g) No, not closed under scalar multiplication.
   (h) No, not closed under addition.
2. (a) Yes ; (b) Yes ; (c) Yes ; (d) No, not closed under scalar multiplication.
   (e) Yes ; (f) No, zero function is not in W. (g) Yes ;
   (h) No, not closed under scalar multiplication.
3. (a) No, not closed under scalar multiplication.
   (b) No, not closed under scalar multiplication.
   (c) No, not closed under scalar multiplication.
   (d) Yes ; (e) Yes
   (f) No, not closed under addition.
   (g) No, not closed under scalar multiplication or zero matrix is not in W.
   (h) Yes.

7. (a), (b) and (d) are linear combinations of $\vec{u}$ and $\vec{v}$.
   (c) Not linear combination.

8. (a) $(-6, -2, 6) = 1\vec{u} + (-1)\vec{v} + 1\vec{w}$.

   (b) $(8, 6, -6) = 0 \cdot \vec{u} + 2\vec{v} + (-2)\vec{w}$.

   (c) $(0, 0, 0) = 0 \cdot \vec{u} + 0 \cdot \vec{v} + 0 \cdot \vec{w}$.

   (d) $(6, -7, 9) = 2\vec{u} + 1\vec{v} + 5\vec{w}$.

9. (a) $7 + 8x - 7x^2 = (-2)\,P_1 + 3P_2$.
   (b) $-8 - 7x + 5x^2 = 1 \cdot P_1 + (-3)\,P_2$.
   (c) $-4 - x - x^2 = (-1)\,P_1 + (-1)\,P_2$.
10. (a), (b) and (c) are all are linear combinations.
11. (a) Yes, (b) Yes, (c) No, (d) No.

12. $\vec{v} = (-6)\,\vec{u_1} + 3\vec{u_2} + 2\vec{u_3}$.

13. $k = -8$.

3.4 Linear Dependence and Independence

Definition : Let V be a vector space and $S = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}\}$ be non-empty set of vectors in V. If there exist scalars $k_1, k_2, \ldots\ldots, k_r$, not all zero, such that

$$k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots\ldots + k_r\vec{v_r} = \vec{0} \qquad \ldots(i)$$

Then the set S is called *linearly dependent* set in V. The set S is called *linearly independent* if S is not dependent. That is, if S is linearly independent, then, equation (i) is satisfied only if $k_1, k_2, \ldots\ldots, k_r$ are all zero. In other words, if S is linearly independent, then the equation (i) has only one solution namely,

$$k_1 = k_2 = \ldots\ldots\ldots = k_r = 0.$$

Note : The empty set ϕ is defined to be linearly independent.

Let us note some of the consequences of the definition of linear dependence and independence, before we workout examples.

Important Observations :

(1) If a subset T of the set S in a vector space V is dependent, then S is also dependent.

For, let $T = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_i}\}$ be the dependent subset of

$S = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_i}, \vec{u_1}, \vec{u_2} \ldots\ldots \vec{u_r}\}$, then there are scalars $k_1, k_2, \ldots\ldots k_r$, not all zero, such that, $k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots + k_r\vec{v_r} = \vec{0}$, since, T is dependent.

Then we have, scalars $k_1, k_2, \ldots\ldots k_r$, not all zero. Such that,

$k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots\ldots + k_r\vec{v_r} + 0.\vec{u_1} + 0.\vec{u_2} + \ldots\ldots + 0.\vec{u_r} = \vec{0}$,

showing that S is dependent.

(2) If a set S contains zero vector, then S is dependent.

(3) Let $S = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}\}$ be a subset of a vector space V. If one vector say $\vec{v_1}$ is scalar multiple of v_2, then S is dependent. For if, $\vec{v_1} = k\vec{v_2}$, $k \neq 0$, otherwise $\vec{v_1} = \vec{0}$, and S is dependent by (2). Thus, we have $0.\vec{v_1} + (-k)\,\vec{v_2} = \vec{0}$ without being $k = 0$, hence, $\{\vec{v_1}, \vec{v_2}\}$ is dependent subset of S, hence, S is also dependent by (1).

(4) (1) can be stated in other words as *'super set of a dependent set is always dependent'*.. This note also implies that any *subset* of an independent set is always independent. Therefore, it is compatible with the fact that the **empty set is defined to be linearly independent, since it is subset of every set**.

SOLVED EXAMPLES

Example 3.39 : The vectors $\overrightarrow{v_1} = (1, -2)$ and $\overrightarrow{v_2} = (-5, 10)$ are dependent in $\mathbb{R}^2$.

Solution : Mere observation shows that

$$\begin{aligned}(+5)\ \overrightarrow{v_1} + \overrightarrow{v_2} &= 5(1, -2) + (-5, 10)\\ &= (5, -10) + (-5, 10)\\ &= (5 + (-5), -10 + 10)\\ &= (0, 0)\end{aligned}$$

That is $5\overrightarrow{v_1} + \overrightarrow{v_2} = 0$, which shows that there exist non-zero scalars $k_1 = 5$, and $k_2 = 1$ such that $k_1\overrightarrow{v_1} + k_2\overrightarrow{v_2} = \overrightarrow{0}$. Therefore, $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ are dependent in $\mathbb{R}^2$.

Example 3.40 : Show that $S = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ is linearly independent in $\mathbb{R}^3$.

Solution : Set $\overrightarrow{v_1} = (1, 1, 1)$, $\overrightarrow{v_2} = (1, 1, 0)$ and $\overrightarrow{v_3} = (1, 0, 0)$. In order to show that S is linearly independent, we have to prove that the equation,

$$k_1\overrightarrow{v_1} + k_2\overrightarrow{v_2} + k_3\overrightarrow{v_3} = 0$$

is satisfied only if $k_1 = k_2 = k_3 = 0$.

The above equation can be written as,

$k_1(1, 1, 1) + k_2(1, 1, 0) + k_3(1, 0, 0) = (0, 0, 0)$.

or $(k_1 + k_2 + k_3, k_1 + k_2, k_1) = (0, 0, 0)$.

Equating the corresponding components we get,

$$\begin{aligned}k_1 + k_2 + k_3 &= 0\\ k_1 + k_2 &= 0\\ k_1 &= 0\end{aligned}$$

Solving these equations for k_1, k_2 and k_3, we have only one solution namely, $k_1 = 0 = k_2 = k_3$.

$\therefore$ S is linearly independent.

Note : Similarly, one can easily check that the set

$$S = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

is linearly independent in $\mathbb{R}^3$.

In general, it can be seen that, the vectors $\overrightarrow{e_1} = (1, 0, 0, 0, \ldots\ldots\ldots\ldots, 0)$, $\overrightarrow{e_2} = (0, 1, 0, \ldots\ldots\ldots, 0)$, …… $\overrightarrow{e_n} = (0, 0, 0, \ldots\ldots\ldots, 1)$ form a linearly independent set of vectors in $\mathbb{R}^n$.

Example 3.41 : Determine whether the set of vectors $\overrightarrow{v_1} = (2, 1, 0, 3)$, $\overrightarrow{v_2} = (3, -1, 5, 2)$ and $\overrightarrow{v_3} = (-1, 0, 2, 1)$ is independent.

Solution : Suppose we have scalars k_1, k_2 and k_3 such that,

$$k_1 \overrightarrow{v_1} + k_2 \overrightarrow{v_2} + k_3 \overrightarrow{v_3} = 0 \quad \ldots (*)$$

or $k_1(2, 1, 0, 3) + k_2 (3, -1, 5, 2) + k_3 (-1, 0, 2, 1) = (0, 0, 0, 0)$

or $(2k_1 + 3k_2 - k_3, k_1 - k_2, 5k_2 + k_3, 3k_1 + 2k_2 + k_3) = (0, 0, 0, 0)$.

Equating the corresponding components we get,

$$\begin{aligned} 2k_1 + 3k_2 - k_3 &= 0 \\ k_1 - k_2 &= 0 \\ 5k_2 + 2k_3 &= 0 \\ 3k_1 + 2k_2 + k_3 &= 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} 5k_2 - k_3 &= 0 \\ 5k_2 + 2k_3 &= 0 \\ 5k_2 + k_3 &= 0 \\ \text{since} \quad k_1 &= k_2 \end{aligned}$$

which shows that, $k_3 = 5k_2$, $k_3 = -\frac{5}{2} k_2$, $k_3 = -5k_2$, which will hold only if $k_2 = 0$, hence $k_3 = 0$ and $k_1 = 0$.

Thus, equation (*) holds only for $k_1 = k_2 = k_3 = 0$.

Therefore, the vectors $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ and $\overrightarrow{v_3}$ are linearly independent in $\mathbb{R}^4$.

Example 3.42 : Determine whether the set of vectors $P_1 = 5$, $P_2 = -6x + x^2$ and $P_3 = (3 - x)^2$ are linearly independent in P_2.

Solution : Suppose we have scalars k_1, k_2 and k_3 such that,

$$k_1 P_1 + k_2 P_2 + k_3 P_3 = 0 \quad \ldots (*)$$

or $k_1(5) + k_2(-6x + x^2) + k_3 (9 - 6x + x^2) = 0$

or $(5.k_1 + 9k_3) + (-6k_2 - 6k_3) x + (k_2 + k_3) x^2 = 0$

Equating the coefficients, we get,

$$\begin{aligned} 5k_1 + 9k_3 &= 0 \\ -6k_2 - 6k_3 &= 0 \quad \ldots (I) \\ \text{and} \quad k_2 + k_3 &= 0 \end{aligned}$$

Solving for k_1, k_2 and k_3, we get,

$$k_1 = -\frac{9}{5} k_3, \; k_2 = -k_3.$$

Thus, the general solution for the system can be written as,

$$k_1 = -\frac{9}{5} t, \; k_2 = -t, \; k_3 = t, \text{ where } t \in \mathbb{R} \quad \ldots (II)$$

which shows that there are infinitely many sets of values of k_1, k_2 and k_3, which will satisfy the equation (*). Therefore the vectors P_1, P_2 and P_3 are not linearly independent.

A particular relation (*) can be obtained by setting $t = -5$, in (II), as

$$+ 9P_1 + (+5) P_2 + (-5) P_3 = 0$$

Example 3.43 : Determine whether the matrices

$A = \begin{pmatrix} 4 & 0 \\ -2 & -2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$ and $C = \begin{pmatrix} 0 & 2 \\ 1 & 4 \end{pmatrix}$ are linearly dependent in $M_{2\times 2}(\mathbb{R})$

Solution : Suppose there are scalars k_1, k_2 and k_3 such that,

$$k_1A + k_2B + k_3C = 0 \qquad \ldots(*)$$

or $\begin{pmatrix} 4k_1 + k_2 & -k_2 + 2k_3 \\ -2k_1 + 2k_2 + k_3 & -2k_1 + 3k_2 + 4k_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Equating the corresponding entries,

$$\begin{aligned} 4k_1 + k_2 &= 0 & k_1 &= -\frac{1}{4}k_2 \\ -k_2 + 2k_3 &= 0 \quad \Rightarrow & k_3 &= \frac{1}{2}k_2 \\ -2k_2 + 2k_2 + k_3 &= 0 & k_3 &= -\frac{5}{2}k_2 \\ -2k_1 + 3k_2 + 4k_3 &= 0 & k_3 &= -\frac{7}{8}k_2 \end{aligned}$$

The last three equations will hold only if $k_2 = 0$, hence, $k_3 = 0$ and $k_1 = 0$. Therefore, the equation (*) holds only if $k_1 = k_2 = k_3 = 0$, so that the matrices A, B and C are linearly independent in $M_{2\times 2}(\mathbb{R})$.

Example 3.44 : Show that the vectors $\overrightarrow{u_1} = e^x$, $\overrightarrow{u_2} = \cos x$, and $\overrightarrow{u_3} = \sin x$ are linearly independent in the function space.

Solution : Suppose

$$ae^x + b\cos x + c\sin x = 0 \qquad \ldots(i)$$

Let $x = 0, \frac{\pi}{2}, \pi$ in (i) then we have three equations,

$$a + b = 0 \qquad \ldots(ii)$$

$$ae^{\pi/2} + c = 0 \qquad \ldots(iii)$$

$$ae^{\pi} + (-b) = 0 \qquad \ldots(iv)$$

From (ii), $a = -b$, putting in (iv), we get $(-b)e^{\pi} - b = 0$ or $b(1 + e^{\pi}) = 0$.

Since, $1 + e^{\pi} \neq 0$, we must have $b = 0$, so that $a = 0$. Using $a = 0$ in (iii) we get $c = 0$.

Thus, the expression (i) holds only if $a = b = c = 0$, showing that $\overrightarrow{u_1}$, $\overrightarrow{u_2}$ and $\overrightarrow{u_3}$ are linearly independent.

Now, it is time to prove some theorems relating to dependent–independent sets.

Theorem 6 : Let S be a set with two or more vectors in a vector space V.

(a) Then S is linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the rest of the vectors in S.

(b) S is linearly dependent if and only if one of the vector in S linear combination of the **preceding vectors**, provided S contains more than one vectors.

Proof : (a) Let $S = \{\vec{u_1}, \vec{u_2} \ldots\ldots, \vec{u_r}\}$ be the set with two or more vectors. If we assume that S is linearly dependent, then there must exist scalars $k_1, k_2, \ldots\ldots, k_r$, not all zero, such that, $k_1\vec{u_1} + k_2\vec{u_2} + \ldots\ldots + k_r\vec{u_r} = \vec{0}$ …(1)

Since $k_1, k_2 \ldots\ldots, k_r$ are not all zero, atleast one member is not zero, so suppose, $k_i \neq 0$, with some i, $1 \leq i \leq k$. Then equation (1) can be written as :

$$k_i \vec{v_i} = (-k_1\vec{v_1}) + (-k_2\vec{v_2}) + \ldots + (-k_{i-1}\vec{v_{i-1}}) + (-k_{i+1}\vec{v_{i+1}}) + \ldots, (-k_r\vec{v_r}) \quad \ldots(2)$$

Since $k_i \neq 0$ is a scalar, multiplying throughout by $\frac{1}{k_i}$ and using vector space axioms, equation (2) takes the form,

$$\vec{v_i} = \left(-\frac{k_1}{k_i}\right)\vec{v_1} + \left(-\frac{k_2}{k_i}\right)\vec{v_2} + \ldots\ldots + \left(-\frac{k_{i-1}}{k_i}\right)\vec{v_{i-1}} + \left(-\frac{k_{i+1}}{k_i}\right)\vec{v_{i+1}} + \ldots\ldots + \left(-\frac{k_r}{k_i}\right)\vec{v_r}.$$

which shows that $\vec{v_i}$ is a linear combination of the rest of the vectors in S.

Conversely, suppose one of the vectors in S is a linear combination of the rest of the vectors in S. In particular, suppose $\vec{v_1}$ is linear combination of the rest of the vectors in S. Then, we have, scalars $k_2, k_3, \ldots\ldots, k_r$ such that

$$\vec{v_1} = k_2\vec{v_2} + k_3\vec{v_3} + \ldots\ldots\ldots + k_r\vec{v_r}$$

Transform $\vec{v_1}$ on R.H.S. and rewriting the above equation we obtain.

$$(-\vec{v_1}) + k_2\vec{v_2} + k_3\vec{v_3} + \ldots\ldots + k_r\vec{v_r} = 0 \text{ or } (-1)\vec{v_1} + k_2\vec{v_2} + \ldots\ldots\ldots + k_r\vec{v_r} = 0$$

This equation holds with scalars $k_1 (= -1), k_2, \ldots\ldots, k_r$, not all zero. Hence by definition the vectors $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_r}$ must be dependent. That is, S is a dependent set. This proves part (a).

(b) Let $S = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}\}$ be the set of vectors in a vector space V.

Suppose S is linearly dependent, then there exist scalars $k_1, k_2, \ldots, k_r$; not all zero such that $k_1\vec{u_1} + k_2\vec{u_2} + \ldots + k_r\vec{u_r} = \vec{0}$. … (i)

Since not $k_1, k_2, \ldots, k_r$ are all zero, let i be the largest suffix such that $k_i \neq 0$; that is; $k_i \neq 0$ and $k_{i+1} = 0 = \ldots = k_r = 0$. Then from (i), we get

$$k_1\vec{u_1} + \ldots + k_i\vec{u_i} = \vec{0} \quad \text{or} \quad -(k_i\vec{u_i}) = k_1\vec{u_1} + k_2\vec{u_2} + \ldots + k_{i-1}\vec{u_{i-1}}$$

Since $k_i \neq 0$, we can write this equation using vector space axioms as :

$$\vec{u_i} = \left(-\frac{k_1}{k_i}\right)\vec{u_1} + \left(-\frac{k_2}{k_i}\right)\vec{u_2} + \ldots + \left(-\frac{k_{i-1}}{k_i}\right)\vec{u_{i-1}}$$

which shows that $\vec{u_i}$ is linear combination of the vector $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_{i-1}}$ - the preceding vectors of $\vec{u_i}$.

Conversely suppose one of the vectors, say $\overrightarrow{u_i}$ is linear combination of the preceding vectors, then for some scalars $k_1, \ldots, k_{i-1}$

$$\overrightarrow{u_i} = k_1\overrightarrow{u_1} + \ldots + k_{i-1}\overrightarrow{u_{i-1}}$$

This can be written as

$$k_1\overrightarrow{u_1} + \ldots + k_{i-1}\overrightarrow{u_{i-1}} + (-1)\overrightarrow{u_i} = 0$$

or $$k_1\overrightarrow{u_1} + \ldots + k_{i-1}\overrightarrow{u_{i-1}} + (-1)\overrightarrow{u_i} + 0 \cdot \overrightarrow{u_{i+1}} + \ldots + 0 \cdot \overrightarrow{u_r} = \overrightarrow{0}$$

$\therefore$ S is linearly dependent.

Theorem 7 : A set with exactly two vectors is linearly dependent if and only if either vector is a scalar multiple of the other.

Proof : Let $S = \{\overrightarrow{u_1}, \overrightarrow{u_2}\}$ be the set with two vectors.

If we assume that one of the vector in S is linear combination of the other, say $\overrightarrow{v_1} = k\overrightarrow{v_2}$, $k \neq 0$, then we have, $v_1 + (-k)\, v_2 = \overrightarrow{0}$, which holds with scalars $k_1(= 1)$, $k_2 (= -k)$, not both zero, hence $S = \{\overrightarrow{v_1}, \overrightarrow{v_2}\}$ is dependent.

Conversely, if we assume that S is linearly dependent, then there exist scalars k_1, k_2, not both zero, such that,

$$k_1\overrightarrow{v_1} + k_2\overrightarrow{v_2} = 0$$

or $$k_1\overrightarrow{v_1} = -k_2\overrightarrow{v_2}$$

or $$\overrightarrow{v_1} = \left(-\frac{k_2}{k_1}\right)\overrightarrow{v_2}$$

$\because k_1 \neq 0$, which shows that v_1 is scalar multiple of the vector v_2.

Note :

(1) The theorem (6) is the negation of the statement : "A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other."

(2) In $\mathbb{R}^3$, a set of three vectors is linearly independent if and only if, the vectors do not lie in the same plane when they are placed with their initial points at the origin.

(3) Vectors in $\mathbb{R}^2$, $\mathbb{R}^3$ or $\mathbb{R}^n$ can be thought of as vectors with their initial points at the origin.

(4) In $\mathbb{R}^3$, a set of three vectors is linearly independent if and only if, the vectors do not lie in the same plane when with their initial points at the origin.

Theorem 8 : A homogeneous system $AX = 0$ of m equations in n unknowns always has a non-trivial soltuion if $m < n$, that is, if the number of unknowns exceeds the number of equations.

Proof : Let R be the reduced row echelon form of A. Then the homogeneous system AX = 0 and RX = 0 are equivlent. If r is the number of non-zero rows of R, then $r \leq m$. If $m < n$ (hypothesis), we conclude that $r < n$. This means, we are solving r equations in n unknowns and can solve for r unknowns (leading variables) in terms of the remaining n – r unknowns, the later being free variables to take on any real numbers. Thus, by letting one of these n – r unknowns be non-zero we obtain a non-trival solution to RX = 0 and thus to AX = 0.

Note : The alternative form of above theorem is : If A is $m \times n$ matrix and AX = 0 has only the trival solution, then $m \geq n$.

Theorem 9 : Let $S = \{\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots\ldots, \overrightarrow{u_t}\}$ be set of vectors in $\mathbb{R}^n$. If $t > n$, then S is linearly dependent.

Proof : Let us write vectors, $\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots\ldots, \overrightarrow{u_t}$ in terms of their components :

$$\overrightarrow{u_1} = (u_{11}, u_{12}, \ldots\ldots\ldots, u_{1n})$$

$$\overrightarrow{u_2} = (u_{21}, u_{22}, \ldots\ldots\ldots, u_{2n})$$

.
.
.
.

$$\overrightarrow{u_t} = (u_{t1}, u_{t2}, \ldots\ldots\ldots, u_{tn}),$$

where, all u_{ij} are real numbers.

Let us consider the equation,

$$k_1\overrightarrow{u_1} + k_2\overrightarrow{u_2} + \ldots\ldots + k_t\overrightarrow{u_t} = 0 \qquad \ldots(1)$$

Writing this equation in terms of the components of $\overrightarrow{u_1}$, $\overrightarrow{u_2}$, ……, $\overrightarrow{u_t}$ and using closure axioms of $\mathbb{R}^n$, we obtain :

$$(k_1 u_{11} + k_2 u_{21} + \ldots\ldots + k_t u_{t1}, k_1 u_{12} + k_2 u_{22} + \ldots\ldots + k_t u_{t2}, \ldots\ldots, k_1 u_{1n}$$
$$+ k_2 u_{2n} + \ldots\ldots + k_t u_{tn}) = (0, 0, \ldots\ldots, 0) \qquad \ldots(2)$$

The right hand side is an ordered n-tupple of zeros.

Equating the components in equation (2) we obtain.

$$u_{11} k_1 + u_{21} k_2 + \ldots\ldots + u_{t1} k_t = 0$$
$$u_{12} k_1 + u_{22} k_2 + \ldots\ldots + u_{t2} k_t = 0$$
$$\begin{array}{cccc} . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \end{array}$$
$$u_{1n} k_1 + u_{2n} k_2 + \ldots\ldots + u_{tn} k_t = 0$$

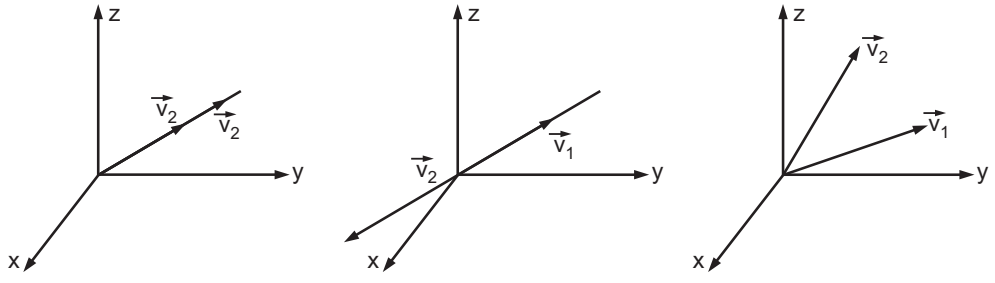
This is a homogeneous system of n-linear equations in t unknowns $k_1, k_2, \ldots\ldots, k_t$. We know that, *'A homogeneous system of linear equations having more unknowns than the number of equations has a non-trivial solution.*' Since t > n, the number of unknowns is greater than number of equations in present case, it has a non-trivial solution. That is, there are scalars $k_1, k_2, \ldots\ldots, k_t$, not all zero, which satisfy the equation (1). Therefore,

$S = \{\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots\ldots, \overrightarrow{u_t}\}$ is linearly dependent set.

Remark 1 : *Above theorem implies in particular that a set in $\mathbb{R}^2$ with more than two vectors is linearly dependent and a set in $\mathbb{R}^3$ with more than three vectors is linearly dependent.*

Geometric interpretation of linear independence :

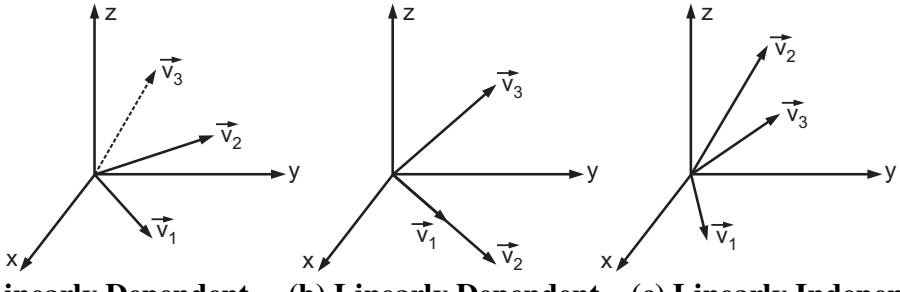
1. In $\mathbb{R}^2$ or $\mathbb{R}^3$, a set of two vectors is linearly independent if and only if the vectors **do not lie on the same line**, when they are placed with their initial points at the origin. This result follows from theorem (9) that two vectors are linearly independent if and only if neither vector is a scalar multiple of the other. Geometrically, this is equivalent to stating that the vectors do not lie on the same line, when they are positioned with their initial points at the origin. See the following figures :



(a) Linearly Dependent **(b) Linearly Dependent** **(c) Linearly Independent**

Fig. 3.1

2. In $\mathbb{R}^3$, a set of three vectors is linearly independent if and only if the vectors **do not lie on the same plane** when they are placed with their initial points at the origin. See the following figure.



(a) Linearly Dependent **(b) Linearly Dependent** **(c) Linearly Independent**

Fig. 3.2

Note : The theorem (9) is the theoretical generalization of the geometric interpretation in $\mathbb{R}^2$ or $\mathbb{R}^3$ as given above.

Linear dependence of functions :

Linear dependence or independence of function is difficult to decide. In example (3.44) above we used trigonometric values or some times we can use some identities that we know. For example, the functions $f_1 = \cos^2 x$, $f_2 = \sin^2 x$ and $f_3 = 3$. From linearly dependent set, since the equation $3f_1 + 3f_2 - f_3 = 3\cos^2 x + 3\sin^2 x - 3 = 3(\cos^2 x + \sin^2 x) - 3 = 0$. In general this may not be the case, a Polish-French mathematician and philosopher J. M. Wronski established a result relating to linear dependence of the functions, which posses continuous derivatives upto some order.

Definition : Let $f_1 = f_1(x)$, $f_2 = f_2(x)$, …, $f_n = f_n(x)$ are $(n-1)$ – times differentiable functions on the interval [a, b], then the determinant.

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \dots & f_n(x) \\ f_1'(x) & f_2'(x) & \dots & f_n'(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \dots & f_n^{(n-1)}(x) \end{vmatrix}$$

is called the **Wronskian** of the functions $f_1, \dots, f_n$.

SOLVED EXAMPLE

Example 3.45 : Let $f_1 = x^2$ and $f_2 = \cos x$, then the Wronskian of f_1 and f_2 is

$$W(x) = \begin{vmatrix} x^2 & \cos x \\ 2x & -\sin x \end{vmatrix}$$

$$= -x^2 \sin x - 2x \cos x$$

Note that : $$W\left(\frac{\pi}{2}\right) = -\frac{\pi^2}{4}\sin\frac{\pi}{2} - \frac{2\pi}{2}\cos\frac{\pi}{2} = -\frac{\pi^2}{4} \neq 0$$

The Wronskian of the functions is useful to determine linear dependence of the functions.

Theorem 10 : Let $f_1, f_2, \dots, f_n$ be the functions having $(x-1)$ continuous derivatives on the interval (a, b) and let the Wronskian of these functions is not identically zero on (a, b), then these functions from a linearly dependent set of vectors in $C^{(n-1)}$ (a, b).

Proof : Let us suppose, for the moment, that $f_1, f_2, \dots, f_n$ are linearly dependent set of vectors in the vector space $C^{(n-1)}$ (a, b). Then there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_n$, not all zero, such that

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + \dots + \alpha_n f_n(x) = 0$$

For all $x \in$ (a, b). Combining this equation with the equations obtained by $(n-1)$ successive differentiation we obtain :

$$\alpha_1 f_1(x) + \alpha_2 f_2(x) + \dots\dots\dots + \alpha_n f_n(n) = 0$$

$$\alpha_1 f_1'(x) + \alpha_2 f_2'(x) + \dots\dots\dots\dots + \alpha_n f_n'(x) = 0$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\alpha_1 f_1^{(n-1)}(x) + \alpha_2 f_2^{(n-1)}(x) + \dots + \alpha_n f_n^{(n-1)}(x) = 0$$

Thus, the linear dependence of $f_1, f_2, \ldots, f_n$ implies that the system of linear equations.

$$\begin{bmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \ldots & f_n^{(n-1)}(x) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

has a non-trivial solution for every $x \in (a, b)$. This implies in turn that for every $x \in (a, b)$, the coefficient matrix is **not invertible**, or, equivalently that its determinant (the Wronskian) **is zero** for every $x \in (a, b)$.

Remark : From theorem (9) it follows that, if the Wronskian of the functions $f_1, f_2, \ldots, f_n$ is **not identically zero** on (a, b), then the functions $f_1, f_2, \ldots, f_n$ must be linearly independent vectors in $c^{(n-1)}[(a, b), \mathbb{R}]$.

For example, in example (3.45) as $W\left(\frac{\pi}{2}\right) \neq 0$, we assert that $f_1 = x^2$, $f_2 = \cos x$ are linearly independent on $(-\pi, \pi)$.

SOLVED EXAMPLE

Example 3.46 : Use the Wronskian to show that the set of vectors e^x, xe^x and x^2e^x is linearly independent.

Solution : We have the Wronskian of e^x, xe^x, x^2e^x is

$$W(x) = \begin{vmatrix} e^x & xe^x & x^2e^x \\ e^x & e^x + xe^x & 2xe^x + x^2e^x \\ e^x & 2e^x + xe^x & 2e^x + 2xe^x + 2xe^x + x^2e^x \end{vmatrix}$$

Using the properties of determinants, we have

$$W(x) = \begin{vmatrix} e^x & 0 & 0 \\ e^x & e^x & 2xe^x \\ e^x & 2e^x & 2e^x + 2xe^x + 2xe^x \end{vmatrix}$$

$$= e^{3x} \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 2x \\ 1 & 2 & 2 + 4x\, e^x \end{vmatrix}$$

$$= e^{3x} [(2 + 4xe^x) - 4x]$$

$$= e^{3x} [2 + 4x (e^x - 1)]$$

so we see that for any $x \in \mathbb{R}$, $W(x) \neq 0$, therefore the vectors e^x, xe^x and x^2e^x are linearly independent.

Exercise 3.3

1. Show that the vectors $\vec{u_1} = (2, -1, 0, 3)$; $\vec{u_2} = (1, 2, 5, -1)$ and $\vec{u_3} = (7, -1, 5, 8)$ are linearly dependent.
2. Show that the polynomials $P_1 = 1 - x$, $P_2 = 5 + 3x - 2x^2$ and $P_3 = 1 + 3x - x^2$ form a linearly dependent set in P_2.
3. Determine whether the vectors $\vec{u_1} = (1, -2, 3)$, $\vec{u_2} = (5, 6, -1)$ and $\vec{u_3} = (3, 2, 1)$ form a linearly dependent set in $\mathbb{R}^3$.
4. Which of the following sets of vectors in $\mathbb{R}^3$ are linearly dependent ?
   (a) (2, 1, 3), (–1, 6, 5)
   (b) (1, 1, 0), (1, 0, 1), (0, 1, 1)
   (c) (1, 2, –3), (1, –3, 2), (2, –1, 5)
   (d) (–1, 6, 5), (0, 0, 0), (1, –3, 2).
   (e) (1, –3, 7), (2, –3, 7), (3, – 1, –4), (1, 1, 1).
5. For which real values of λ do the following vectors form linearly dependent set in $\mathbb{R}^3$?
   $\vec{u_1} = \left(-\frac{1}{2}, -\frac{1}{2}, \lambda\right)$; $\vec{u_2} = \left(\lambda, -\frac{1}{2}, -\frac{1}{2}\right)$, $\vec{u_3} = \left(-\frac{1}{2}, \lambda, -\frac{1}{2}\right)$.
6. Let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be independent vectors. Show that $\vec{u} + \vec{v}$, $\vec{u} - \vec{v}$ and $\vec{u} - 2\vec{v} + \vec{w}$ are also linearly independent.
7. Show that in a vector space V, every single non-zero vector is linearly independent.
8. If $S = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_r}\}$ is linearly dependent set in a vector space V and $\vec{u} \in V$, then show that $S' = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_r}, \vec{u}\}$ is also dependent.
9. If $S = \{\vec{u_1}, \vec{u_2}, \ldots\ldots \vec{u_r}\}$ is linearly independent set in a vector space V, and $S' = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_r}, \vec{v}\}$ is dependent, show that $\vec{v}$ is a linear combination of vectors in S.
10. Show that the vectors $\vec{v_1} = e^{a_1 x}$, $\vec{v_2} = e^{a_2 x}$,……, $\vec{v_n} = e^{a_n x}$ in a function space are linearly independent, where $a_1, a_2, \ldots\ldots, a_n$ are distinct real numbers.
11. Show that $f_1 = 1$, $f_2 = e^x$ and $f_3 = e^{2x}$ from a linearly independent set of vectors in $c^2(-\infty, \infty)$.
12. Which of the following set of vectors in F [(a, b), $\mathbb{R}$] are linearly dependent.
   (a) $3\cos^2 x, 4\sin^2 x, 12$.
   (b) $x, \cos x$.
   (c) $\cos 2x, \sin^2 x, \cos^2 x$.
   (d) $(3 - x)^2, x^2 - 6x, 5$.
   (e) $1, \cos x, \cos 2x$.

13. Use Wronskian to show that the following vectors are linearly independent.
    (a) e^x, e^{2x}, e^{3x}.
    (b) sin x, cos x, x sin x.
    (c) $1, x, x^2$.
    (d) $1, 1 + x, 1 + x^2, 1 + x^3$.

Answers (3.3)

1. $3\vec{u_1} + \vec{u_2} - \vec{u_3} = \vec{0}$.
2. $3p_1 - p_2 + 2p_3 = 0$
3. Yes
4. (a) Not linearly dependent.
   (b) Not linearly dependent.
   (c) Not linearly dependent.
   (d) Yes
   (e) Yes
5. $\lambda = -\frac{1}{2}$; $\lambda = 1$.

10. Use mathematical induction on n. For n = 1, the result is obvious. Suppose for scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$.

    $\alpha_1 a^{a_1 x} + \alpha_2 e^{a_2 x} + \ldots\ldots \alpha_k e^{a_k x} + \alpha_{k+1}{}^{a_{k+1} x} = 0$.

    Let $a_1 = \max \{a_1, a_2, \ldots\ldots, a_k, a_{k+1}\}$, w.l.g. and multiply the above equation by $e^{-a_1 x}$ and let $x \to \infty$, show that $\alpha_1 \cdot 1 + 0 + \ldots + 0 = 0$ or $\alpha_1 = 0$. Then induction applies.

12. (a) Linearly dependent as : $4 (3 \cos^2 x) + 3 (4 \sin^2 x) + (-1)\, 12 = 0$.
    (b) Linearly independent.
    (c) Linearly dependent : $\because \sin^2 x - \cos^2 x - 1 \cos 2x = 0$.
    (d) Linearly dependent as : $(-1)(3 - x)^2 + 1 (x^2 - 6x) + \left(-\frac{9}{5}\right) 5 = 0$.
    (e) Linearly independent.

3.5 Basis and Dimension

Definition : If V is any vector space and $B = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_n}\}$ is a set of vectors in V, then B is called a basis for V if the following conditions hold :

(i) B is linearly independent; and,
(ii) B spans V.

Remark : A basis is the vector space generalization of a co-ordinate system in 2-space and 3-space, this will be clear when we shall define co-ordinates of vectors in a vector space, in a little while.

Theorem 10 (a) : If $B = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_n}\}$ is a basis for a vector space V, then every vector $\vec{v}$ in V can be expressed in the form $\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + \ldots\ldots + c_n\vec{v_n}$ in exactly one way, where $c_1, c_2, \ldots, c_n$ are scalars.

Proof : Since B spans V, it follows from the definition of a spanning set that every vector in V is expressible as a linear combination of the vectors in B. Suppose that $\vec{v}$ can be written as,

$$\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + \ldots\ldots\ldots + c_n\vec{v_n}$$

and also as,

$$\vec{v} = k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots + k_n\vec{v_n}$$

This shows that,

$$\vec{0} = \vec{v} - \vec{v} = (c_1\vec{v_1} + c_2\vec{v_2} + \ldots\ldots + c_n\vec{v_n}) - (k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots + k_n\vec{v_n})$$

$$= (c_1 - k_1)\vec{v_1} + (c_2 - k_2)\vec{v_2} + \ldots\ldots + (c_n - k_n)\vec{v_n},$$

(by using vector space axioms)

Thus,

$$(c_1 - k_1)\vec{v_1} + (c_2 - k_2)\vec{v_2} + \ldots\ldots + (c_n - k_n)\vec{v_n} = \vec{0}$$

Since the left side of this equation is linear combination of vectors in B, the linear independence of B implies that $c_1 - k_1 = 0$, $c_2 - k_2 = 0$, ……, $c_n - k_n = 0$, that is, $c_1 = k_1$, $c_2 = k_2$, ……, $c_n = k_n$.

Thus, the two expressions for $\vec{v}$ are the same.

$\therefore$ $\vec{v}$ can be expressed uniquely.

This theorem enables us to define co-ordinates of any vector in a vector space, which is the generalized form of the co-ordinates of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, when referred to the rectangular co-ordinate system of $\mathbb{R}^2$ or $\mathbb{R}^3$.

Definition : If $B = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_n}\}$ is a basis for a vector space V, and for $\vec{v}$ in V, we have, $\vec{v} = c_1\vec{v_1} + c_2\vec{v_2} + \ldots\ldots + c_n\vec{v_n}$ is the expression for $\vec{v}$ in terms of the basis B, then the scalars $c_1, c_2, \ldots\ldots, c_n$ are uniquely determined by the basis B. (by theorem 8), these scalars $c_1, c_2, \ldots\ldots, c_n$ determined by B for $\vec{v}$, are called the *co-ordinates* of $\vec{v}$

relative to the basis B. The vector $(c_1, c_2, \ldots\ldots, c_n)$ corresponding to these scalars (co-ordinates) is called the co-ordinate vector of $\vec{v}$ relative to B, and it is denoted by

$$(\vec{v})_B = (c_1, c_2, \ldots\ldots, c_n)$$

Note : Co-ordinate of vector depend not only on the basis B but also on the order in which the basis vectors are written.

SOLVED EXAMPLES

Example 3.47 : Let $V = \mathbb{R}^3$ and $B = \{\vec{i}, \vec{j}, \vec{k}\}$ where, $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ and $\vec{k} = (0, 0, 1)$. Then we have shown that B is linearly independent set in $\mathbb{R}^3$. Also B spans $\mathbb{R}^3$, since any vector (a, b, c) in $\mathbb{R}^3$ can be written as,

$$\vec{v} = (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1).$$
$$= a\vec{i} + b\vec{j} + c\vec{k}.$$

Thus the co-ordinates of $\vec{v}$ are a, b, c and the co-ordinate vector is $\vec{v}$ itself.
This basis B is called the ***standard basis*** or ***natural basis*** for $\mathbb{R}^3$.

Note :

(1) From above example, we see that for any $\vec{v} = (a, b, c)$ in $\mathbb{R}^3$, the co-ordinates of $\vec{v}$ relative to the standard basis are a, b, and c so $(\vec{v})_B = (a, b, c)$ which shows that $\vec{v} = (\vec{v})_B$.

(2) Similarly the set $B = \{(1, 0), (0, 1)\}$ is the standard basis for $\mathbb{R}^2$.

Example 3.48 : This is the generalization of the example 1 and note 2 above. The set of vectors $\vec{e_1} = (1, 0, 0, \ldots\ldots, 0)$, $\vec{e_2} = (0, 1, 0, \ldots\ldots, 0)$, ……, $\vec{e_n} = (0, 0, 0, \ldots\ldots, 1)$ are linearly independent, which we have seen earlier. Therefore, the set $B = \{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}\}$ is basis for $\mathbb{R}^n$, because B spans $\mathbb{R}^n$ as any vector $\vec{u} = (\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n})$ in $\mathbb{R}^n$ can be written as,

$$\vec{u} = u_1\vec{e_1} + u_2\vec{e_2} + \ldots\ldots + u_n\vec{e_n} \qquad \ldots(*)$$

This basis B for $\mathbb{R}^n$, is called the **standard basis** or **natural basis** for $\mathbb{R}^n$. It is clear from (*) that the co-ordinates of u relative to the standard basis are $u_1, u_2, \ldots\ldots, u_n$.

So $(\vec{u})_B = (u_1, u_2, \ldots\ldots, u_n)$.

Hence, $(\vec{u})_B = \vec{u}$.

So the vector u and its co-ordinate vector relative to standard basis are the same.

Example 3.49 : Show that the set $B = \{\vec{u_1}, \vec{u_2}\}$ is a basis for $\mathbb{R}^2$, where $\vec{u_1} = (3, 4)$ and $\vec{u_2} = (1, -3)$. Also find co-ordinate vector of the vector (1, 2) relative to this basis.

Solution : To show that B spans $\mathbb{R}^2$, we must show that an arbitrary vector $\vec{v} = (a, b)$ in $\mathbb{R}^2$ can be expressed as a linear combination $\vec{v} = (a, b) = k_1\vec{u_1} + k_2\vec{u_2}$ of vectors in B. Writing right side in terms of components gives

$$\begin{aligned}(a, b) &= k_1(3, 4) + k_2(1, -3)\\ &= (3k_1 + k_2, 4k_1 - 3k_2)\end{aligned}$$

Or on equating the corresponding components ;

$$\left.\begin{aligned}3k_1 + k_2 &= a\\ 4k_1 - 3k_2 &= b\end{aligned}\right\} \quad \ldots(i)$$

Thus, to show that B spans $\mathbb{R}^2$, we must prove that system (∗) has a solution for all choices of a and b, and that is for all choices of v in $\mathbb{R}^2$.

Solving the system (i) for k_1, k_2 we obtain

$$k_1 = \frac{1}{13}(3a + b)$$

and $$k_2 = \frac{4}{13}a - \frac{3}{13}b.$$

showing that the system (∗) has solution for all choices of $\vec{v}$ in $\mathbb{R}^2$. Therefore, B spans $\mathbb{R}^2$.

Now, to prove that B is linearly independent we must show that the only solution of $k_1\vec{u_1} + k_2\vec{u_2} = \vec{0}$ is $k_1 = k_2 = 0$. As above expressing this in terms of components, the verification of independence reduces to the homogeneous system showing that

$$\left.\begin{aligned}3k_1 + k_2 &= 0\\ 4k_1 - 3k_2 &= 0\end{aligned}\right\} \quad \ldots(ii)$$

has only trivial solution. It is easy to solve the system (∗∗) for k_1 and k_2 which gives $k_1 = 0$ and $k_2 = 0$. Thus, B is linearly independent.

Therefore, B spans $\mathbb{R}^2$ and is linearly independent, hence is a basis for $\mathbb{R}^2$.

Example 3.50 : Let $\vec{u_1} = (1, 2, -3)$, $\vec{u_2} = (1, -3, 2)$, and $\vec{u_3} = (2, -1, 5)$ be vectors in $\mathbb{R}^3$. Then show that the set $B = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ is a basis for $\mathbb{R}^3$.

Solution : To show that B spans $\mathbb{R}^3$, we must prove that any arbitrary vector $\vec{u} = (a, b, c)$ in $\mathbb{R}^3$ can be expressed as a linear combination

$$\vec{u} = k_1 u_1 + k_2 u_2 + k_3 u_3$$

of vectors in B. Converting this equation in terms of components we get,

$$\vec{u} = (a, b, c) = k_1 (1, 2, -3) + k_2 (1, -3, 2) + k_3 (2, -1, 5)$$

or $(a, b, c) = (k_1 + k_2 + 2k_3, 2k_1 - 3k_2 - k_3, -3k_1 + 2k_2 + 5k_3)$

or on equating components

$$\left.\begin{aligned} k_1 + k_2 + 2k_3 &= a \\ 2k_1 - 3k_2 - 1k_3 &= b \\ -3k_1 + 2k_2 + 5k_3 &= c \end{aligned}\right\} \quad \ldots(*)$$

Thus, to show that B spans $\mathbb{R}^3$, we must prove that the system (*) has a solution for all choices of $\vec{u} = (a, b, c)$ in $\mathbb{R}^3$.

Also, to prove that B is linearly independent, we must prove that the only solution of

$$k_1\vec{u_1} + k_2\vec{u_2} + k_2\vec{u_3} = 0$$

is $k_1 = k_2 = k_3 = 0$. As above, expressing this equation in terms of components, the verification reduces to showing that the homogeneous system

$$\begin{aligned} k_1 + k_2 + 2k_3 &= 0 \\ 2k_1 - 3k_2 - k_3 &= 0 \\ -3k_1 + 2k_2 + 5k_3 &= 0 \end{aligned} \quad \ldots(**)$$

has only the trivial solution.

Observe that systems (*) and (**) have the same coefficient matrix. Thus by using theorems on solutions of linear systems of equations, we can simultaneously prove that B is linearly independent and spans $\mathbb{R}^3$, by showing that in systems (*) and (**) the matrix of coefficients,

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -3 & -1 \\ -3 & 2 & 5 \end{bmatrix}$$

is invertible; since det (A) = –30 (compute) ; is that, det (A) ≠ 0. Thus, B is a basis for $\mathbb{R}^3$.

Note : The systems in (*) and (**) in above example can be solved by either method of elimination or by Gauss elimination or by Gauss-Jordan elimination. Also matrix inversion method can be applied (if det (A) ≠ 0) since A is square matrix. In present example A is a square marix, and we need to establish the existence of solutions, therefore, it suffices to show that det (A) ≠ 0 to state that B is basis for V, which is the convenient way. When specific solution is required we need to elaborate to find the solution by using either of the methods mentioned above.

Example 3.51 : (a) Find the co-ordinate vector of the vector (1, 2) relative to the basis obtained in example (3.49) for $\mathbb{R}^2$.

(b) Find the coordinate vector of (1, 2, 3) relative to the basis obtained in example (4) above. Also find a vector whose co-ordinate vector relative to this basis is (3, 2, 1).

Solution : (a) We have to find scalars k_1 and k_2, such that,

$\vec{u} = (1, 2) = k_1\vec{u_1} + k_2\vec{u_2}$

or $(1, 2) = k_1(3, 2) + k_2(1, -3)$

or $(1, 2) = (3k_1 + k_2, 4k_1 - 3k_2)$

or on equating the corresponding components.

$$3k_1 + k_2 = 1.$$
$$4k_1 - 3k_2 = 2.$$

Solving this system, we obtain

$$k_1 = \frac{5}{13}$$

and $$k_2 = -\frac{2}{13} \text{ (Verify !)}$$

Thus, $$(\vec{u})_B = \left(\frac{5}{13}, -\frac{2}{13}\right).$$

(b) Here also, we must find scalars k_1, k_2 and k_3 such that

$\vec{u} = (1, 2, 3) = k_1\vec{u_1} + k_2\vec{u_2} + k_3\vec{u_3}$

or $(1, 2, 3) = k_1(1, 2, -3) + k_2(1, -3, 2) + k_3(2, -1, 5)$

or $(1, 2, 3) = (k_1 + k_2 + 2k_3, 2k_1 - 3k_2 - k_3, -3k_1 + 2k_2 + 5k_3)$

or on equating the corresponding components.

$$\left.\begin{aligned} k_1 + k_2 + 2k_3 &= 1 \\ 2k_1 - 3k_2 - k_3 &= 2 \\ -3k_1 + 2k_2 + 5k_3 &= 3 \end{aligned}\right\} \quad \ldots(i)$$

The matrix of coefficients of the system (∗) is,

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -3 & -1 \\ -3 & 2 & 5 \end{bmatrix}$$

and we know that det (A) = –30. The inverse of A by calculating co-factor matrix can obtained as :

$$A^{-1} = \frac{1}{\det(A)} \text{adj } A = -\frac{1}{30}\begin{bmatrix} -13 & -1 & 5 \\ -7 & 11 & 5 \\ -5 & -5 & -5 \end{bmatrix}$$

$$\therefore \quad A^{-1} = \begin{bmatrix} \frac{13}{30} & \frac{1}{30} & -\frac{5}{30} \\ \frac{7}{30} & -\frac{11}{30} & -\frac{5}{30} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

Hence by matrix inversion method,

$$\begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = A^{-1} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{13}{30} & \frac{1}{30} & -\frac{5}{30} \\ \frac{7}{30} & -\frac{11}{30} & -\frac{5}{30} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

Thus, the co-ordinates of $\vec{u} = (1, 2, 3)$ relative to the basis B are 0, –1 and 1. Hence,

$$(\vec{u})_B = (0, -1, 1)$$

Now, to prove another part, by definition of co-ordinate vector $(u)_s$, we have

$$\vec{u} = k_1\vec{u_1} + k_2\vec{u_2} + k_3\vec{u_3}$$

or $\vec{u} = (3)\,\vec{u_1} + (2)\,\vec{u_2} + (1)\,\vec{u_3}$, … here $k_1 = 3$, $k_2 = 2$ and $k_3 = 1$ are given

$= (3)\,(1, 2, -3) + (2)\,(1, -3, 2) + 1\,(2, -1, 5)$

$= (7, -1, 10)$

Thus the vector (7, –1, 10) has co-ordinate vector (3, 2, 1) relative to the basis B.

Example 3.52 : Show that $B = \{M_1, M_2, M_3, M_4\}$ where, $M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is a basis for the vector space $M_{2\times 2}(R)$ of 2×2 matrices.

(b) Find the co-ordinate vector of $M = \begin{bmatrix} 2 & -3 \\ 4 & -5 \end{bmatrix}$ relative to the above basis.

Solution : (a) Any arbitrary matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ can be written as,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = aM_1 + bM_2 + cM_3 + dM_4,$$

thus B spans $M_{2\times 2}(\mathbb{R})$

On the other hand any equation of the form,

$aM_1 + bM_2 + cM_3 + dM_4 = 0$

i.e. $a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

i.e. $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ has unique solution a = b = c = d = 0.

This shows that B is linearly independent. This basis B is called the standard basis for $M_{2 \times 2}(\mathbb{R})$. More generally, the standard basis for $M_{m \times n}(\mathbb{R})$ consists of the mn distinct matrices with single 1 and zeros for the remaining entries.

(b) Since the matrix M can be written as

$$M = \begin{bmatrix} 2 & -3 \\ 4 & -5 \end{bmatrix} = 2M_1 + (-3)\, M_2 + 4M_3 + (-5)\, M_4$$

the co-ordinate vector of M relative to the basis B is (2, –3, 4, –5). Thus,

$$(M)_B = (2, -3, 4, -5).$$

Example 3.53 : (a) Show that the set $B = \{1, x, x^2, \ldots\ldots, x^n\}$ is a basis for the space P_n of polynomials of degree $\leq n$.

(b) Find the co-ordinate vector of the polynomial

$$P = a_0 + a_1x + a_2x^2 + a_3x^3$$

relative to the basis

$$B = \{x^3, x^2, x, 1\} \text{ for } P_3.$$

Solution : (a) Any polynomial P is expressed as $P = a_0 + a_1x + \ldots\ldots + a_n \mid x_n$, which is clearly a linear combination of the vectors in B. Therefore, B spans P.

On the other hand, if any linear combination of vectors in B equals zero, that is,

$a_0 + a_1x + a_2x^2 + \ldots\ldots + a_nx^n = 0$, for all x,

then we must have $a_0 = a_1 = \ldots\ldots = a_n = 0$.

So that B is linearly independent and thus, B is a basis for P_n. This basis B is called **standard** basis or **natural basis** for P_n.

Note : The set $B' = \{x^n, x^{n-1}, \ldots, x, 1\}$ is also a basis for P_n (verify!)}.

(b) Since we can write P as

$$P = a_3x^3 + a_2x^2 + a_1x + a_0 1$$

a linear combination of vectors in $B = \{x^3, x^2, x, 1\}$, the co-ordinates of P relative to this basis B are a_3, a_2, a_1 and a_0. Therefore, the co-ordinate vector of P is (a_3, a_2, a_1, a_0). That is,

$$(P)_B = (a_3, a_2, a_1, a_0)$$

Note : If $B = \{\vec{v}_1, \vec{v}_2, \ldots\ldots, \vec{v}_t\}$ is linearly independent set of vectors in a vector space V, then B is basis for L (B) or lin (B), since B spans L (B) and is linearly independent.

Result 2 : Let $S = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be a set of non-zero vectors in a vector space V and let W = L(S). Then some subset of S is a basis for W.

Proof : First, if S is linearly independent and as S spans W hence we conclude that S is a basis for W.

Now suppose S is linearly dependent, then

$$k_1\vec{u_1} + k_2\vec{u_2} + \ldots + k_n\vec{u_n} = \vec{0} \quad \ldots (i)$$

where $k_1, k_2, \ldots, k_n$ are scalars, not all zero. Then one of the vectors in S is a linear combination of the remaining vectors of S, say $\vec{u_i}$. Then, one we can drop this vector from S and still retain the spanning set of W; that is if $S_1 = \{\vec{u_1}, \ldots, \vec{u_{i-1}}, u_{i+1}, \ldots, \vec{u_n}\}$, then $W = L(S_1)$.

Again if S_1 is linearly independent, then S_1 is a basis for W. If S_1 is linearly dependent then we repeat the above argument to S_1 in place of S and get a new set S_2 which spans W. Continuing, since S is a finite set, we will find a subset T of S such that T is linearly independent and spans W. The set T, then is a basis for W.

Note :

(1) The above theorem gives a simple procedure for finding a subset T of a set S so that T is a basis for L(S). But this is tedius procedure at each step we have to solve the homogeneous system of inear equations to find $k_1, k_2, \ldots, k_n$ and droping a vector with non-zero constant k_i,

(2) We have the alternative procedure if $V = \mathbb{R}^n$.

Let S be subset of $V = \mathbb{R}^n$ and let W = L(S).

Let $S = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be the set of non-zero vectors in V. The procedure for finding a subset of S that is a basis for W = L(S) as follows :

Step 1 : From the equation

$$k_1\vec{u_1} + k_2\vec{u_2} + \ldots + k_n\vec{u_n} = \vec{0} \quad \ldots (*)$$

Construct the augmented matrix associated with the homogeneous system (*) and transform it to reduced row echelon form.

Step 2 : The row vectors corresponding to the columns containing the leading ones form a basis for W = L(S).

We shall use this procedure in later part of the chapter.

Definition : A non-zero vector space V is called *finite dimensional* if it contains a finite set of vectors $\{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n}\}$ that forms a basis for V. If no such set exists, V is called infinite dimensional.

For example : The vector spaces $\mathbb{R}^2$, $\mathbb{R}^3$, or $\mathbb{R}^n$, P_n and $M_{m \times m}(\mathbb{R})$ are finite dimensional vector spaces, as we have seen in above examples that each of these vector spaces have bases with finite number of vectors.

Note : We regard zero vector space to be finite dimensional, though it has no linearly independent set, consequently no basis. The function spaces F([a, b], $\mathbb{R}$), C[a, b), D[a, b] and R[a, b] are infinite dimensional.

Theorem 11 : If $B = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n}\}$ is a basis for a vector space V, then every set with more than n vectors from V is linearly dependent.

Proof : Let $B' = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_m}\}$ be any set of m vectors in V, where m > n. We have to show that B' is linearly dependent. Since B is basis for V, each v_i can be expressed as a linear combination of the vectors in B, say,

$$\left.\begin{aligned}
\vec{v_1} &= a_{11}\vec{u_1} + a_{12}\vec{u_2} + \ldots\ldots\ldots + a_{1n}\vec{u_n} \\
\vec{v_2} &= a_{21}\vec{u_1} + a_{22}\vec{u_2} + \ldots\ldots\ldots + a_{2n}\vec{u_n} \\
&\vdots \\
\vec{v_m} &= a_{m1}\vec{u_1} + a_{m2}\vec{u_2} + \ldots\ldots\ldots + a_{mn}\vec{u_n}
\end{aligned}\right\} \quad \ldots (*)$$

To show that B' is linearly dependent, we must prove that there exist scalars $k_1, k_2, \ldots, k_m$, not all zero such that

$$k_1\vec{v_1} + k_2\vec{v_2} + \ldots\ldots + k_m\vec{v_m} = \vec{0} \quad \ldots(I)$$

Using the equations in (*), we can rewrite (I) as :

$$(k_1a_{11} + k_2a_{12} + \ldots\ldots + k_ma_{1m})\vec{u_1} + (k_1a_{12} + k_2a_{22} + \ldots\ldots + k_ma_{m2})\vec{u_2}$$

$$+ \ldots\ldots + (k_1a_{1n} + k_2a_{2n} + \ldots\ldots + k_ma_{mn})\vec{v_n} = 0$$

Since B is linearly independent, the long expression which equals to zero, gives

$$\left.\begin{aligned}
a_{11}k_1 + a_{12}k_2 + \ldots\ldots + a_{1m}k_m &= 0 \\
a_{12}k_1 + a_{22}k_2 + \ldots\ldots + a_{m2}k_m &= 0 \\
&\vdots \\
a_{1n}k_1\, a_{2n}k_2 + \ldots\ldots + a_{mn}\, k_m &= 0
\end{aligned}\right\} \quad \ldots(**)$$

Thus, the problem of proving that B' is linearly dependent set reduces to showing that there are scalars $k_1, k_2, \ldots\ldots, k_m$, not all zero, which satisfy the homogeneous system (**) in m-unknowns and in n equations. Since m > n, that is, the number of unknowns is more than the number of equations (**) has a non-trivial solution, hence, the proof follows.

Note : From theorem (11), we see that if $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ is a basis for a vector space V, then the number of vectors in linearly independent set cannot exceed n.

Theorem 12 : If V is a finite dimensional vector space, then any two bases for V have the same number of vectors.

Proof : Suppose $B = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n}\}$ and $B' \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_m}\}$ are two bases for V. We have to show that n = m.

Since B is basis and B' is linearly independent, by theorem (10), it follows that $m \leq n$. Similarly, since B' is basis and B is linearly independent set, we must have $n \leq m$. Therefore, m = n.

Note :

(1) The standard basis for $\mathbb{R}^n$ contains n vectors, therefore, every basis for $\mathbb{R}^n$, contains exactly n vectors. Therefore, any set with more than n vectors is linearly dependent in $\mathbb{R}^n$.

(2) The standard basis for P_n contains (n + 1) vectors, thus, every basis for P_n contains n + 1 vectors.

The theorem above enables us to define the concept of dimension of a vector space.

Definition : If V is finite-dimensional vector space, then the *dimension of V* is defined as the number of vectors in a basis for V. In addition, we define the **zero vector** space to have **dimension zero**.

Note :

(1) The dimension of zero vector space is zero, which is appropriate as we have already defined that zero vector space is spanned by an empty set, and empty set contains no elements.

(2) The dimension of a vector space V is denoted by dim V.

Thus, we have, from above illustrated examples, that,

$\dim \mathbb{R}^2 = 2$, $\dim \mathbb{R}^3 = 3$, $\dim \mathbb{R}^n = n$, $\dim P_n = n + 1$ and $\dim M_{m \times n}(\mathbb{R}) = mn$.

Theorem 13 : Let V be an n-dimensional vector space then,

(a) Any set with n linearly independent vectors in V is a basis for V.

(b) Any set with n vectors which spans V is a basis for V.

Proof : (a) Let $B = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n}\}$ be any linearly independent set with n-vectors in V. To prove that B is a basis for V, we need to show that B spans V. Let $\vec{v}$ be any vector in V which is not in B. Then by theorem, the set $B' = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n}, \vec{v}\}$ is

linearly dependent as it contains more than n-vectors. Therefore, there exist scalars, k_1, k_2,……, k_n and k, not all zero such that,

$$k_1\vec{u_1} + k_2\vec{u_2} + \ldots\ldots + k_n\vec{u_n} + k\vec{v} = \vec{0} \qquad \ldots(I)$$

If k = 0, the equation (I) reduces to

$$k_1\vec{u_1} + k_2\vec{u_2} + \ldots\ldots + k_n\vec{u_n} = \vec{0},$$

and since B is linearly independent, this must imply that $k_1 = k_2 = \ldots = k_n = 0$, which contradicts to the assumption, that B' is dependent set. Therefore, $k \neq 0$. Then from equation (I) we can write,

$$\vec{v} = \left(-\frac{k_1}{k}\right)\vec{u_1} + \left(-\frac{k_2}{k}\right)\vec{u_2} + \ldots\ldots + \left(-\frac{k_n}{k}\right)\vec{u_n}, \quad \because k \neq 0.$$

This shows that $\vec{v}$ is a linear combination of vectors in B, and as $\vec{v}$ is arbitrary, it follows that B spans V. Therefore, B must be a basis for V.

(b) Let $B = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_n}\}$ be a set of n vectors in V and dim V = n such that B spans V. To show that B is a basis for V, we need to show that B is linearly independent. Suppose B is not linearly independent, that is, it is linearly dependent then by theorem (9) (a), at least one of the vectors in B is linear combination of the remaining vectors in B. This tells us that we can retain a set B' with (n – 1) vectors from B by deleting the vector which is linear combination of the other vectors in B and that B' spans V. But this contradiction as dim V = n, V is can not be spanned by less than n-vectors. Hence B must be linearly independent, so B is a basis for V.

Note : From above theorem it follows that if dimension of a vector space is n and if B is any set with n vectors in V, which is either linearly independent or which spans V, then B is a basis for V.

Theorem 14 : Let V be a vector space with dim V = n.

(a) If $S = \{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_t}\}$ is a linearly independent set of vectors in V, and t < n, then S can be enlarged to a basis for V; that is, there are vectors $\vec{u}_{t+1}, \vec{u}_{t+2}, \ldots\ldots \vec{u_n}$ such that the set $\{\vec{u_1}, \vec{u_2}, \ldots\ldots, \vec{u_t}, \vec{u}_{t+1}, \ldots\ldots, \vec{u_n}\}$ is a basis for V.

(b) If W is a subspace of V, then dim W ≤ dim V ; moreover, dim W = dim V; if and only if W = V.

Proof : (a) Let $\vec{v}$ be any non-zero vector in V. Consider the set $S' = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_t}, \vec{v}\}$. If S' is linearly dependent, there exist scalars $k_1, k_2, \ldots\ldots, k_t$ and k, not all zero, such that,

$$k_1\vec{u_1} + k_2\vec{u_2} + \ldots\ldots\ldots + k_t\vec{u_t} + k\vec{v} = 0 \qquad \ldots(I)$$

If k = 0, the linearly independence of S implies that $k_1 = k_2 = \ldots\ldots = k_t = 0$, hence k cannot be zero. Then from equation (I), we can express :

$$\vec{v} = \left(-\frac{k_1}{k}\right)\vec{u}_1 + \left(-\frac{k_2}{k}\right)\vec{u}_2 + \ldots\ldots + \left(-\frac{k_t}{k}\right)\vec{u}_t$$

which shows that $\vec{v}$ is linear combination of vectors in S. Since $\vec{v}$ is arbitrary, it follows that S spans V, hence S is a basis for V. But this is contradiction to the fact that dim V = n as t < n. Therefore, S' must be linearly independent. If t + 1 = n, then by theorem 16 (a), S' is a basis for V. If t + 1 < n, then proceed as before with S' in place of S and obtain a set S" with t + 2 vectors which is linearly independent. If t + 2 = n, then again by theorem 16 (a) S" will be a basis for V.

If t + 2 < n, then proceed as before with S" in place of S'. Since n is finite, we obtain vectors $\vec{u}_{t+1}, \vec{u}_{t+2}, \ldots\ldots, \vec{u}_n$ such that the set $\{\vec{u}_1, \vec{u}_2, \ldots\ldots, \vec{u}_t, \vec{u}_{t+1}, \ldots\ldots, \vec{u}_n\}$ is linearly independent. So by theorem 16 (a), it is a basis for V.

(b) Since W is subspace of V, any linearly independent set in W is linearly independent in V. Therefore, if $S = \{\vec{u}_1, \vec{u}_2, \ldots\ldots, \vec{u}_t\}$ is basis for W, then it is linearly independent set in V also. So by theorem (14) $t \leq n$, since dim (V) = n.

Thus, dim (W) ≤ dim (V).

Now, if dim (W) = dim (V) and $S = \{\vec{u}_1, \vec{u}_2, \ldots\ldots, \vec{u}_n\}$ is basis for W, then it is linearly independent as it is basis for W, then it is linearly independent in V also, hence by theorem 16 (a), it is basis for V also. Therefore, W = V.

Conversely, if W = V, then clearly any basis for V is a basis for W.

Therefore, dim (W) = dim (V).

SOLVED EXAMPLES

Example 3.54 : If $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis for a vector space V, show that $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is also a basis, where $\vec{u}_1 = \vec{v}_1$, $\vec{u}_2 = \vec{v}_1 + \vec{v}_2$ and $\vec{u}_3 = \vec{v}_1 + \vec{v}_2 + \vec{v}_3$.

Solution : To show that $= \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is basis for V, by theorem 14, it suffices to show that either S is linearly independent or spans V. We opt to show that S is linearly independent. Suppose we have scalars k_1, k_2 and k_3 such that

$$k_1\vec{u}_1 + k_2\vec{u}_2 + k_3\vec{u}_3 = \vec{0}$$

or writing $\vec{u}_1$, $\vec{u}_2$ and $\vec{u}_3$ in terms of $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$,

$$k_1\vec{v}_1 + k_2(\vec{v}_1 + \vec{v}_2) + k_3(\vec{v}_1 + \vec{v}_2 + \vec{v}_3) = \vec{0}\ .$$

or $(k_1 + k_2 + k_3)\vec{v}_1 + (k_2 + k_3)\vec{v}_2 + k_3\vec{v}_3 = \vec{0}$.

Since v_1, v_2 and v_3 are linearly independent, this implies that,

$$\begin{aligned} k_1 + k_2 + k_3 &= 0 \\ k_2 + k_3 &= 0 \\ k_3 &= 0 \end{aligned}$$

Solving this we get, $k_1 = k_2 = k_3 = 0$.

Therefore, $S = \{\overrightarrow{u_1}, \overrightarrow{u_2}, \overrightarrow{u_3}\}$ is linearly independent.

Example 3.55 : Determine whether the set $B = \{(3, 1, -4), (2, 5, 6), (1, 4, 8)\}$ is basis for $\mathbb{R}^3$.

Solution : Since, any basis of $\mathbb{R}^3$ contains three vectors, by theorem 16 (a), it suffices to show that either S is linearly independent or S spans $\mathbb{R}^3$.

To show that S is linearly independent, suppose we have scalars k_1, k_2 and k_3 such that,

$$k_1 (3, 1, -4) + k_2 (2, 5, 6) + k_3 (1, 4, 8) = (0, 0, 0) \quad \ldots(i)$$

$$\therefore \quad (3k_1 + 2k_2 + k_3, k_1 + 5k_2 + 4k_3, -4k_1 + 6k_2 + 8k_3) = (0, 0, 0)$$

or on equating corresponding components,

$$\begin{aligned} 3k_1 + 2k_2 + k_3 &= 0 \\ k_1 + 5k_2 + 4k_3 &= 0 \quad \ldots(ii) \\ k_1 + 6k_2 + 8k_3 &= 0 \end{aligned}$$

Thus, the problem of showing S is linearly independent reduces to showing that the system (ii) has only trivial solution. And by theorem (b) and (d) stated in beginning of this section, it suffices to show that the matrix of coefficients

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 5 & 4 \\ -4 & 6 & 8 \end{bmatrix}$$

has non-zero determinant.

$$\begin{aligned} \det(A) &= 3(40 - 24) - 2(8 + 16) + 1(6 + 20). \\ &= 3 \times (16) - 2 \times (-24) - 26. \\ &= 26 \neq 0. \end{aligned}$$

Therefore, S is linearly independent and hence, is a basis for $\mathbb{R}^3$.

Note : The matrix A of the coefficients of the system (∗) consists of rows which are the vectors in S. So this suggests that in $\mathbb{R}^3$ (or $\mathbb{R}^2$ or $\mathbb{R}^4$), if we want to show that set of three (or two or four) vectors is a basis for $\mathbb{R}^3$ (corresponding for $\mathbb{R}^2$ or $\mathbb{R}^4$), just find the determinant of a matrix A whose rows are the vectors of the given set. If $\det(A) \neq 0$, then S will be linearly independent and hence be a basis for V.

If $\det(A) = 0$, then S cannot be linearly independent, so that it cannot be a basis.

For example, the set of vectors (2, 3) and (6, 9) is not basis for $\mathbb{R}^2$, as

$$\det(A) = \det\begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix} = 0$$

whereas, the set of vectors (1, –1) and (2, 2) forms a basis for $\mathbb{R}^2$ since,

$$\det(A) = \det\begin{bmatrix} 1 & -1 \\ 2 & 2 \end{bmatrix} = 4 \neq 0.$$

Example 3.56 : Determine the dimension and a basis for the solution space of the system

$$x_1 - 4x_2 + 3x_3 - x_4 = 0$$
$$2x_1 - x_2 + 5x_3 - 3x_4 = 0.$$

Solution : The augmented matrix of given system is

$$\begin{bmatrix} 1 & -4 & 3 & -4 & 0 \\ 2 & -1 & 5 & -3 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbb{R}_2 - 2\mathbb{R}_1} \begin{bmatrix} 1 & -4 & 3 & -4 & 0 \\ 0 & 7 & -1 & -1 & 0 \end{bmatrix}$$

i.e. $x_1 - 4x_2 + 3x_3 - x_4 = 0$ … (1)

$$7x_2 - x_3 - x_4 = 0 \quad \ldots (2)$$

Set $x_3 = 4$, $x_s = t$.

$\therefore$ $7x_2 = s + t \therefore x_2 = \dfrac{s + t}{7}$

$\therefore$ From (1), $x_1 = \dfrac{-178 + 11t}{7}$

$\therefore$ The solution set of given system is

$$\left\{\left(\frac{-17s + 11t}{7}\right), \frac{s + t}{7}, s, t \mid s, t \in R\right\}$$

$$= \left\{s\left(-\frac{17}{7}, \frac{1}{7}, 1, 0\right) + t\left(\frac{11}{7}, \frac{1}{7}, 0, 1\right)\right\}$$

$$= \{s(-17, 1, 7, 0) + t(11, 1, 0, 7)\}$$

It can be seen that the vectors (–17, 1, 7, 0) and (11, 1, 0, 7) are linearly independent. Hence the dimension of solution space is 2 and its basis is {(17, 1, 7, 0), (11, 1, 0, 7)}.

Exercise 3.4

1. Explain in each case why the following set of vectors is not basis for the indicated vector spaces. (Solve this problem by just inspection).

   (a) $\overrightarrow{u_1} = (2, 1)$, $\overrightarrow{u_2} = (4, 3)$, $\overrightarrow{u_3} = (5, 2)$ for $\mathbb{R}^2$

   (b) $\overrightarrow{u_1} = (1, 1, 1)$, $\overrightarrow{u_2} = (1, 1, 0)$ for $\mathbb{R}^3$

(c) $\overrightarrow{u_1} = (-1, 1, 1)$, $\overrightarrow{u_2} = (0, 0, 0)$, $\overrightarrow{u_3} = (0, 1, -1)$ for $\mathbb{R}^3$

(d) $P_1 = 1 - x + x^2$, $P_2 = 1 + x$ for P_2.

(e) $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix}$, $D = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$

$E = \begin{bmatrix} 5 & -1 \\ 3 & 5 \end{bmatrix}$, for $M_{2\times 2}(\mathbb{R})$

(f) $\overrightarrow{u_1} = (-1, 1, 1)$, $\overrightarrow{u_2} = (0, 1, -1)$, $\overrightarrow{u_3} = \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$

2. Determine which of the following sets of vectors is basis for $\mathbb{R}^2$. (Give reasons).

(a) (1, 2), (1, 0) (b) (0, 0), (–1, 2) (c) (1, 1), (–1,0)

(d) $(-\frac{1}{2}, 1)$, (– 1, 2) (e) (1, 2), (2, 1), (1, 1)

3. Determine which of the following sets of vectors are bases for $\mathbb{R}^3$. (Give reasons).

(a) (5, 5, 5), (4, 4, 0), (3, 0, 0)

(b) (1, 2, –3), (1, –3, 2), (2, –1, 5).

(c) (0, 0, 1), (0, 1, 1), (1, 1, 1)

(d) (1, –3, 7), (2, 0, –6), (3, –1, –1), (2, 4, 5)

(e) (1, –1, 1), (3, 1, –1)

(f) (1, 1, –1), (1, –2, 3), (2, –1, 2)

4. Show that the following set of vectors is a basis for $M_{2\times 2}(\mathbb{R})$

$$\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -8 \\ -12 & -4 \end{bmatrix}, \begin{bmatrix} 3 & 6 \\ 3 & -6 \end{bmatrix}$$

5. Find the co-ordinate vector of $\begin{bmatrix} 4 & 3 \\ -11 & -8 \end{bmatrix}$ relative to the basis given in exercise (4).

6. Find the co-ordinate vector of A relative to the basis $B = \{A_1, A_2, A_3, A_4\}$.

$A = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix}$, $A_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $A_4 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$.

7. Find the co-ordinate vector of $\overrightarrow{u}$ relative to the basis $S = \{\overrightarrow{u_1}, \overrightarrow{u_2}\}$ for $\mathbb{R}^2$

(a) $\overrightarrow{u_1} = (0, 2)$, $\overrightarrow{u_2} = (1, 1)$, $\overrightarrow{u} = (a, b)$

(b) $\overrightarrow{u_1} = (2, -3)$, $\overrightarrow{u_2} = (6, 0)$, $\overrightarrow{u} = (10, -15)$

(c) $\overrightarrow{u_1} = (0, 1)$, $\overrightarrow{u_2} = (1, -1)$, $\overrightarrow{u} = (1, 2)$.

8. Find the co-ordinate vector of $\vec{u}$ relative to the basis B = $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$.

   (a) $\vec{u} = (3, 2, 1)$, $\vec{v_1} = (1, 1, 1)$, $\vec{v_2} = (0, 1, 1)$, $\vec{v_3} = (0, 0, 1)$.

   (b) $\vec{u} = (1, 1, 1)$, $\vec{v_1} = (1, 2, 3)$, $\vec{v_2} = (-4, 5, 6)$, $\vec{v_3} = (7, -8, 9)$.

   (c) $\vec{u} = (2, -1, 3)$, $\vec{v_1} = (3, 0, 0)$, $\vec{v_2} = (2, 2, 0)$, $\vec{v_3} = (1, 1, 1)$.

In exercise 9 – 13 determine the dimension and basis for the solution space of the system.

9. $x_1 + 2x_2 - 4x_3 + 3x_4 - x_5 = 0$
   $x_1 + 2x_2 - 2x_3 + 2x_4 + x_5 = 0$
   $2x_1 + 4x_2 - 2x_3 + 3x_4 + 4x_5 = 0$

10. $x_1 + 2x_2 + 7x_3 = 0$
    $-2x_1 + x_2 - 4x_3 = 0$
    $x_1 - x_2 + x_3 = 0$.

11. $x_1 + 3x_2 + 2x_3 = 0$.
    $x_1 + 5x_2 + x_3 = 0$
    $3x_1 + 5x_2 + 8x_3 = 0$

12. $x_1 - 2x_2 + 7x_3 = 0$
    $2x_1 + 3x_2 - 2x_3 = 0$
    $2x_1 - x_2 + x_3 = 0$

13. $x_1 + 4x_2 + 2x_3 = 0$
    $2x_1 + x_2 + 5x_3 = 0$

14. Let W = $\{(x, y, z, w) \mid x + y + z = 0\}$. Find the basis and dimension of the subspace W of $\mathbb{R}^4$.

15. Let W = $\{(x, y, z, w) \mid y + z = 0, x = 2w\}$ be a subspace of $\mathbb{R}^4$. Find the basis and dimension of W.

Answers (3.4)

1. (a) A basis for $\mathbb{R}^2$ contains two vectors.
   (b) A basis for $\mathbb{R}^3$ contains three vectors.
   (c) The set is dependent; since zero vector is included.
   (d) P_2 has basis with three elements.
   (e) A basis for $M_{2\times 2}(\mathbb{R})$ has four vectors.
   (f) The set is dependent as $\vec{u_1} + 0.\vec{u_2} + 2\vec{u_3} = 0$.

2. (a) Yes, (b) No, (c) Yes , (d) No, since $\vec{u_1} = -\frac{1}{2}\vec{u_2}$. (e) No.

3. (a), (b), (c) Yes, since the sets are independent with three elements.
   (d) No, since basis of $\mathbb{R}^3$ has three vectors.
   (e) No, since basis of $\mathbb{R}^3$ has three vectors.
   (f) No, since $\vec{u_1} + \vec{u_2} - \vec{u_3} = 0$.

5. (1, 1, 1, 1).

6. $(A)_B = (-1, 1, -1, 3)$.

7. (a) $(\overrightarrow{u})_S = (\frac{1}{2}(b - a), a)$, (b) $(\overrightarrow{u})_S = (5, 0)$ (c) $(\overrightarrow{u})_S = (3, 1)$

8. (a) $(3, -1, -1)$, (b) $\left(\frac{7}{10}, -\frac{2}{15}, -\frac{1}{30}\right)$, (c) $(1, -2, 3)$.

9. Basis : $\{(-2, 1, 0, 0, 0), (-1, 0, \frac{1}{2}, 1, 0), (-3, 0, -1, 0, 1)\}$; dimension = 3.

10. Basis : $\{(1, -2, 1), (0, 5, -3)\}$; dimension = 2.

11. Basis : $\{(7, -1, -2)\}$; dimension = 1,

12. Dimension = 0.

13. Basis : $\{(18, -1, -7)\}$; dimension = 1.

14. Basis : $\{(-1, 1, 0, 0), (-1, 0, 1, 0), (0, 0, 0, 1)\}$; dimension W = 3.

15. Basis : $\{(2, 0, 0, 1), (0, -1, 1, 0)\}$; dimension W = 2.

3.6 Row Space, Column Space, Null Space, Rank and Nullity

In this section we shall state and prove some theorems which will provide us with a deeper understanding of the relationships between the solutions of a linear system and properties of its coefficient matrix. These results will also help us to ease the calculations of the problems such as finding the basis and dimension of the space spanned by the given set of vectors. However, the proof of the theorems in this section are given to understand the results in proper way, but they are not expected in view point of examination.

Definition : Let,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

be an $m \times n$ matrix. The vectors

$$\left.\begin{aligned} R_1 &= (a_{11}, a_{12}, \dots\dots, a_{1n}) \\ R_2 &= (a_{21}, a_{22}, \dots\dots, a_{2n}) \\ &\vdots \\ R_m &= (a_{m1}, a_{m2}, \dots\dots, a_{mn}) \end{aligned}\right\}$$

formed from the rows of A are called the *row vectors* of A, and the vectors.

$$C_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ . \\ . \\ . \\ . \\ a_{m1} \end{bmatrix}, C_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ . \\ . \\ . \\ . \\ a_{m2} \end{bmatrix}, \ldots\ldots\ldots, C_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ . \\ . \\ . \\ . \\ a_{mn} \end{bmatrix}$$

from the columns of A are called the *column vectors* of A.

For example if $A = \begin{bmatrix} -1 & 2 & 0 \\ 5 & 1 & -3 \end{bmatrix}$

The row vectors of A are,

$R_1 = (-1, 2, 0)$, $R_2 = (5, 1, -3)$ and the column vectors of A are,

$$C_1 = \begin{bmatrix} -1 \\ 5 \end{bmatrix}, C_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ and } C_3 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

If A is an $m \times n$ matrix with row vectors $R_1, R_2, \ldots\ldots, R_m$ then it can be denoted by

$$A = \begin{bmatrix} R_1 \\ - \\ R_2 \\ - \\ : \\ : \\ - \\ R_m \end{bmatrix}$$

If A is an $m \times n$ matrix with column vectors $C_1, C_2, \ldots\ldots, C_n$, then it can be denoted by

$$A = [\, C_1 : C_2 : \ldots\ldots\ldots : C_n \,]$$

Definition : If A is an $m \times n$ matrix, then the subspace of $\mathbb{R}^n$ spanned by the row vectors of A is called the *row space* of A, and the subspace of $\mathbb{R}^m$ spanned by the column vectors of A is called the *column space* of A. The solution space of the homogeneous system of equations $AX = 0$, which is a subspace of $\mathbb{R}^n$ is called the *null space of A*.

Definition : If A is $m \times n$ matrix, the dimension of the row space of A is called row-rank of A and the dimension of the column space is called the column rank of A. The dimension of the null space of A is called nullity of A.

Theorem 15 : Let $L = \{R_1, R_2, \ldots, R_m\}$ be the row vectors of a matrix of order $m \times n$. (Clearly the row vectors are vectors in $\mathbb{R}^n$). If e is an elementary row operation, then e(L) is linearly independent if and only if L is linearly independent.

Proof : If $e = R_{ij}$, the interchange of i^{th} row vector and j^{th} row vector, or if $e = kR_i$, multiplying the i^{th} row vector by a non-zero vector k, then the result is obvious. Let $e = R_i + kR_j$, then $e(L) = \{R_1, R_2, \ldots, R_i + kR_j, \ldots R_j, \ldots, R_m\}$. Let L be linearly independent, and suppose that

$$\alpha_1 R_1 + \alpha_2 R_2 + \ldots + \alpha_i (R_i + kR_j) + \ldots + \alpha_j R_j + \ldots + \alpha_m R_m = 0$$

where, $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$. Then this implies

$$\alpha_1 R_1 + \ldots + \alpha_i R_i + \ldots + (\alpha_j + k\alpha_i) R_i + \ldots + \alpha_m R_m = 0.$$

Hence, $\alpha_1 + \ldots = \alpha_j + k\alpha_i = \ldots = \alpha_m = 0$, since L is linearly independent; which further gives $\alpha_1 = \ldots = \alpha_m = 0$.

Therefore, e(L) is a linearly independent set. Conversely if e(L) = L' is linearly independent, then $L = e^{-1}(L')$, where e^{-1} is inverse of e is also an elementary row operation. By previous result itself, L is linearly independent.

Theorem 16 : Let $L = \{X_1, X_2, \ldots, X_m\}$ be the row vectors of a matrix of order $m \times n$. If e is an elementary row operation on L, then L and e(L) have the same (maximal) number of linearly independent elements.

Proof : If $e = R_{ij}$ or $e = kR_i$ type elementary row operation, then the result is obvious. So let $e = R_i + kR_j$. Suppose $e(L) = \{y_1, y_2, \ldots, y_m\}$ where $y_i = X_i + kR_j$, and $y_i = X_i$ if $i \neq p$. Let r be the number of (maximal) linearly independent elements in L. We shall prove that e(L) does not have more than r linearly independent elements. Consider any subset M of e(L) containing r + 1 elements, say

$$M = \{y_{p_1}, y_{p_2}, \ldots y_{p_{r+1}}\}$$

Writing $I = [p_1, p_2, \ldots, p_r, p_{r+1}\}$. If $i, j \in I$ or if $i \in I$, it follows from theorem (13) that M is linearly dependent. So, let us assume that $i \in I$, say $i = p_s$, but $j \in I$. Now the subset

$\{X_{p_1}, \ldots, X_{p_{s-1}}, X_i, X_{p_{s+1}}, \ldots, X_{p_{r+1}}\}$ is linearly dependent and hence there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_{r+1}$, not all zero, such that

$$\alpha_1 X_{p_1} + \ldots + \alpha_s X_i + \ldots + \alpha_{r+1} X_{p_{r+1}} = 0 \quad \ldots \text{(i)}$$

Similarly, the subset $\{X_{p_1}, \ldots, X_{p_{s-1}}, X_j, X_{p_{s+1}}, \ldots, X_{p_{r+1}}\}$ is linearly dependent and hence there exist scalars $\beta_1, \ldots \beta_{r+1}$, not all zero, such that

$$\beta_1 X_{p_1} + \ldots + \beta_s X_j + \ldots + \beta_{r+1} X_{p_{r+1}} = 0 \quad \ldots \text{(ii)}$$

If $\alpha_s = 0$ or $\beta_s = 0$, then M is linearly dependent. If $\alpha_s \neq 0$ and $\beta_s \neq 0$, it follows from (1) and (ii) that

$$\frac{1}{\alpha_s}(\alpha_1 X_{p_1} + \ldots + \alpha_s X_i + \ldots + \alpha_{r+1} X_{p_{r+1}}) + \frac{k}{\beta_s}(\beta_1 X_{p_1} + \ldots + \beta_s X_j + \ldots + \beta_{r+1} X_{p_{r+1}}) = 0 \quad \ldots \text{(iii)}$$

But (iii) implies that M is linearly dependent.

Thus, every subset of e(L) having r + 1 elements is linearly dependent and therefore e(L) does not have more than r linearly independent elements. Since e is invertible, it follows that the number of linearly independent elements in L cannot exceed the number of linearly independent elements in e(L). Hence, L and e(L) have the same number of linearly independent elements.

Theorem 17 : Let A be an $m \times n$ matrix with real entries. Let $A^1, A^2, \ldots, A^n$ denote the column vectors of A. Suppose e is an elementary row operation and e(A) = B. If $B^1, B^2, \ldots, B^n$ denote column vectors of B, then

$$\alpha_1 A^1 + \alpha_2 A^2 + \ldots + \alpha_n A^n = 0$$

holds if and only if,

$$\alpha_1 B^1 + \alpha_2 B^2 + \ldots + \alpha_n B^n = 0, \text{ where } \alpha_1, \alpha_2, \ldots, \alpha_n \text{ belong to } \mathbb{R}.$$

Proof : Suppose $\alpha_1 A^1 + \alpha_2 A^2 + \ldots + \alpha_n A^n = 0$.

This can be written as

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ : \\ \alpha_n \end{bmatrix} = 0$$

Performing the elementary row operation e on both sides and using the theorem which that $e(AB) = e(A) \cdot B$, we get,

$$e(A) \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ : \\ \alpha_n \end{bmatrix} = 0$$

This implies

$$B \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ : \\ \alpha_n \end{bmatrix} = 0, \text{ since e (A) = B}$$

Hence, $\alpha_1 B^1 + \alpha_2 B^2 + \ldots + \alpha_n B^n = 0$

The converse follows from the fact that the elementary operation is invertible.

Remark : Evidently, the same result holds for any subset of $B^1, B^2, \ldots, B^n$ and the corresponding subset of $A^1, A^2, \ldots, A^n$.

Hence, the maximal number of linearly independent columns of B is equal to the maximal number of linearly independent columns of A.

A similar result holds for an elementary column operation on A.

Thus, from above theorem and remark, we have proved a theorem.

Theorem 18 : If A is an $m \times n$ matrix and e is an elementary row (or column) operation on A, then A and e(A) have the same row rank and the same column rank.

Definition : Let A and B be $m \times n$ matrices with real entries. If B can be obtained by applying successively a number of elementary row operations on A, then A is said to be row-equivalent to B, written $A \overset{R}{\sim} B$.

From theorem (19) and the definition, we have,

Theorem 19 : Let A and B be $m \times n$ matrices. If $A \overset{R}{\sim} B$, then A and B have the same row rank and the same column rank.

Proof : Left for exercise.

We prove one more theorem.

Theorem 20 : If A is $m \times n$ row reduced echelon matrix with r non-zero rows, then the row rank of A is r.

Proof : Suppose $R_1, R_2, \ldots, R_m$ are row vectors of A and suppose $R_1, R_2, \ldots, R_r$ are first non-zero row vectors of A. Suppose further that

$$\alpha_1 R_1 + \alpha_2 R_2 + \ldots + \alpha_r R_r = 0;\ \alpha_1, \alpha_2, \ldots, \alpha_r \in R.$$

Let the first non-zero entry in R_i occurs in k_i^{th} column, $i = 1, 2, \ldots, r$. Equating the k_i^{th} components of the n-tuples on two sides of the above equation, we get $\alpha_i = 0;\ i = 1, 2, \ldots, r$. Hence, the rows $R_1, R_2, \ldots, R_r$ are linearly independent. Therefore, using theorem (20) rank $A = r$.

From above discussion it also follows that

Theorem 21 : If a matrix is in reduced row echelon form, then the column vectors that contain leading 1's form a basis for the column space of the matrix.

Proof : Left for exercise.

SOLVED EXAMPLES

Example 3.57 : Find a basis for the null space of

$$A = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$$

Solution : The null space of A is the solution space of the homogeneous system,

$$\left.\begin{aligned} x_1 + 4x_2 + 5x_3 + 6x_4 + 9x_5 &= 0 \\ 3x_1 - 2x_2 + x_3 + 4x_4 - x_5 &= 0 \\ -x_1 - x_3 - 2x_4 - x_5 &= 0 \\ 2x_1 + 3x_2 + 5x_3 + 7x_4 + 8x_5 &= 0 \end{aligned}\right\} \quad \ldots(*)$$

To obtain the general solution of the system (*), we reduce the coefficient matrix A to reduced row-echelon form. For this apply $R_2 + (-3)\,R_1$, $R_3 + R_1$, $R_4 + (-2)\,R_1$ on A. Then we get,

$$\begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 0 & -14 & -14 & -14 & -28 \\ 0 & 4 & 4 & 4 & 4 \\ 0 & -5 & -5 & -5 & -5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By applying $\left[-\frac{1}{14}\right] R_2$ and then $R_3 + (-4)R_2$ and $R_4 + (5)\,R_2$ on previous matrix

$$= \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By applying $R_1 + (-4)R_2$

The last matrix is in reduced row-echelon form of A. The system of linear equations corresponding to last matrix is,

$$x_1 + x_3 + 2x_4 + x_5 = 0$$
$$x_2 + x_3 + x_4 + 2x_5 = 0$$

or
$$x_1 = -x_3 - 2x_4 - x_5$$
$$x_2 = -x_3 - x_4 - 2x_5$$

Here, x_3, x_4 and x_5 are free variables. The general solution of this system is given by,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -r - 2s - t \\ -r - s - 2t \\ r \\ s \\ t \end{bmatrix}$$

$$= \begin{bmatrix} -r \\ -r \\ r \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -2s \\ -s \\ 0 \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} -t \\ -2t \\ 0 \\ 0 \\ t \end{bmatrix} = r \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

This shows that the basis for this system is

B = { (–1, –1, 1, 0, 0), (–2, –1, 0, 1, 0), (–1, –2, 0, 0, 1)}.

and dimension is 3. Therefore, the basis for the null space of A is given by B.

Example 3.58 : Find a basis for the subspace of $\mathbb{R}^4$ spanned by the vectors (1, 1, –4, –3), (2, 0, 2, –2) and (2, –1, 3, 2),

Solution : The space spanned by the given vectors is the row space of the matrix

$$\begin{bmatrix} 1 & 1 & -4 & -3 \\ 2 & 0 & 2 & -2 \\ 2 & -1 & 3 & 2 \end{bmatrix}$$

Applying $R_2 + (-2)\, R_1$ and $R_3 + (-2)\, R_1$ on this matrix we get,

$$\begin{bmatrix} 1 & 1 & -4 & -3 \\ 0 & -2 & 10 & 4 \\ 0 & -3 & 11 & 8 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & -4 & -3 \\ 0 & 1 & -5 & -2 \\ 0 & 0 & -4 & 20 \end{bmatrix}$$

By applying $\left(-\frac{1}{2}\right) R_2$ and then $R_3 + (3)\, R_2$.

$$\begin{bmatrix} 1 & 1 & -4 & -3 \\ 0 & 1 & -5 & -2 \\ 0 & 0 & 1 & -5 \end{bmatrix}$$

This is in row-echelon form.

The non-zero row vectors in this matrix are $\overrightarrow{v_1} = (1, 1, -4, -3)$, $\overrightarrow{v_2} = (0, 1, -5, -2)$ and $\overrightarrow{v_3} = (0, 0, 1, -5)$. These vectors form a basis for the row space of (by theorem 16) the matrix, consequently form a basis for the space spanned by the given vectors.

Example 3.59 : Find a basis for the subspace of $\mathbb{R}^3$ spanned by the vectors (1, 2, –1), (4, 1, 3), (5, 3, 2) and (2, 0, 2).

Solution : The subspace spanned by the given vectors is the row space of the matrix

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 4 & 1 & 3 \\ 5 & 3 & 2 \\ 2 & 0 & 2 \end{bmatrix}$$

formed by the given vectors as rows of A. We shall reduce this matrix to its row-echelon form. For this first apply $R_2 + (-4)R_1$, $R_3 + (-5)\, R_1$ and $R_4 + (-2)\, R_1$ to obtain

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -7 & 7 \\ 0 & -7 & 7 \\ 0 & -4 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

By applying $\left(-\frac{1}{7}\right) R_2$, then apply $R_3 + 7R_2$ and $R_4 + 4R_2$

This matrix is in row-echelon form. The non-zero rows of this matrix are the vectors (1, 2, –1), and (0, 1, –1). These vectors form a basis for the row space of the matrix A consequently form a basis for the space spanned by the given vectors.

Example 3.60 : Find the basis for column space of A consisting entirely column vectors of A. Where,

$$A = \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix}$$

Solution : Let us obtain the reduced row-echelon form of A. Apply $R_2 + 2R_1$, $R_4 + (-3)\, R_1$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & -1 & 1 & 3 & 2 \\ 0 & -3 & 3 & 4 & 1 \\ 0 & 0 & 0 & -13 & -13 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & 1 & -1 & -3 & -2 \\ 0 & 0 & 0 & -5 & -5 \\ 0 & 0 & 0 & -13 & -13 \end{bmatrix}$$

By applying $(-1)\, R_2$, then apply $R_3 + 2R_2$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & 1 & -1 & -3 & -2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By applying $\left(-\frac{1}{5}\right)R_3$ and then apply $R_4 + 13R_3$

This matrix is in row-echelon form. We proceed further to obtain its reduced row-echelon form. For this apply $R_2 + (3)R_2$ and $R_1 + (-2)\, R_3$.

$$\begin{bmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 1 & -1 & 0 & +1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$R = \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix R is in reduced row-echelon form.

It is easy to check that the column vectors

$$C_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, C_2' = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } C_3' = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R; thus the corresponding column vectors of A, namely,

$$C_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 3 \end{bmatrix}, C_2 = \begin{bmatrix} 2 \\ -5 \\ -3 \\ 6 \end{bmatrix} \text{ and } C_3 = \begin{bmatrix} 2 \\ -1 \\ 4 \\ -7 \end{bmatrix}$$

form a basis form the column space of A.

Example 3.61 : Find a subset of the vectors that forms a basis for the space spanned by the vectors $\overrightarrow{v_1} = (1, -2, 0, 3)$, $\overrightarrow{v_2} = (2, -4, 0, 6)$, $\overrightarrow{v_3} = (-1, 1, 2, 0)$ and $\overrightarrow{v_4} = (0, -1, 2, 3)$; then express each vector that is not in the basis as a linear combination of the basis vectors.

Solution : Let us construct a matrix A that has $\overrightarrow{v_1}$, $\overrightarrow{v_2}$, $\overrightarrow{v_3}$ and $\overrightarrow{v_4}$ as its column vectors.

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -4 & 1 & -1 \\ 0 & 0 & 2 & 2 \\ 3 & 6 & 0 & 3 \end{bmatrix}$$

$$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$$

$$v_1 \quad v_2 \quad v_3 \quad v_4$$

Let us find a basis for the column space of A. For this we reduce A to reduced row-echelon form. For this we apply $R_2 + (2)R_1$ and $R_4 + (-3)\ R_1$ to yield

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix}$$

Further apply $(-1)\ R_2$ and then $R_3 + (-2)\ R_2$, $R_4 + (-3)\ R_2$ and $R_1 + R_2$.

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$$

$$w_1 \quad w_2 \quad w_3 \quad w_4$$

This last matrix is in reduced row-echelon form. Let us denote its column vectors by $\vec{w_1}, \vec{w_2}, \vec{w_3}$ and $\vec{w_4}$. The leading 1's occur in columns 1 and 3 so that by theorem 16 $\{\vec{w_1}, \vec{w_3}\}$ is a basis for the column space of the last matrix and consequently $\{\vec{v_1}, \vec{v_3}\}$ is a basis for the column space of A; which is the required subset of given vectors and forms a basis for the space spanned by $\vec{v_1}, \vec{v_2}, \vec{v_3}$ and $\vec{v_4}$.

Now, we shall express $\vec{w_2}$ and $\vec{w_4}$ as linear combination of $\vec{w_1}$ and $\vec{w_3}$. The simplest way of doing this is to express $\vec{w_2}$ and $\vec{w_4}$ in terms of basis vectors with smaller subscripts.

It is simple to note that $\vec{w_2} = 2\vec{w_1}$ and $\vec{w_4} = \vec{w_1} + \vec{w_3}$. The corresponding relationships in A are,

$$\vec{v_2} = 2\vec{v_1}$$

and $$\vec{v_4} = \vec{v_1} + \vec{v_3}$$

Example 3.62 : Find a basis for the row space of

$$A = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

consisting entirely of row vectors from A.

Solution : We will transpose A, thereby converting the row space of A into the column space of A^t, then we will use the method of column space of A^t; and then we will transpose again to convert column vectors back to row vectors. Transposing A, we get,

$$A^t = \begin{bmatrix} 1 & 2 & 0 & 2 \\ -2 & -5 & 5 & 6 \\ 0 & -3 & 15 & 18 \\ 0 & -2 & 10 & 8 \\ 3 & 6 & 0 & 6 \end{bmatrix}$$

Reducing this to row-echelon form, we obtain

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & -5 & -10 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first, second and the fourth column contain the leading 1's so the corresponding column vectors in A^t form a basis for the column space of A^t.

$$C_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \\ 3 \end{bmatrix}; C_2 = \begin{bmatrix} 2 \\ -5 \\ -3 \\ -2 \\ 6 \end{bmatrix} \text{ and } C_4 = \begin{bmatrix} 2 \\ 6 \\ 18 \\ 8 \\ 6 \end{bmatrix}$$

Transposing again and adjusting the notations yields the basis vectors.

$R_1 = (1, -2, 0, 0, 3)$,

$R_2 = (2, -5, -3, -2, 6)$

and $R_4 = (2, 6, 18, 8, 6)$ for the row space of A.

Algorithm : Given a set of vectors $S = \{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_k}\}$ in $\mathbb{R}^n$. The following procedure produces a subset of those vectors that is a basis for L(S) and express those vectors of S that are not in the basis as linear combinations of the basis vectors.

Step 1 : Form the matrix A having $\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_k}$ as its column vectors.

Step 2 : Reduce matrix A to its reduced row-echelon form R, and let $\vec{w_1}, \vec{w_2}, \ldots\ldots, \vec{w_k}$ be the column vectors of r.

Step 3 : Identify the columns that contain the leading 1's in R. The corresponding column vectors of A are the basis vectors for L(S).

Step 4 : Express each column vector of R that does not contain a leading 1 as a linear combination of proceeding column vectors that do contain leading 1's. This yields a set of dependency equations involving the column vectors of R. The corresponding equations for the column vectors of A express the vectors not in the basis as linear combinations of the basis vectors.

Definition : Given a matrix A, the common dimension of the row space and column space of a matrix A is called the *rank* of A and is denoted by *rank (A)*; the dimension of the null space of A is called the *nullity* of A and is denoted by *nullity (A).*

Example 3.63 : Find the rank and nullity of the matrix

$$A = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$$

Solution : We reduce A to its reduced row-echelon form. For this apply $R_2 + (-3)\, R_1$, $R_3 + R_1$ and $R_4 + (-2)\, R_1$ to obtain

$$\begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 0 & -14 & -14 & -14 & -14 \\ 0 & 4 & 4 & 4 & 8 \\ 0 & -5 & -5 & -5 & -10 \end{bmatrix}$$

Apply $\left(-\frac{1}{14}\right) R_2$ and then $R_3 + (-4)\, R_2$, $R_4 + (5)\, R_2$ to obtain

$$\begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Apply $R_1 + (-4)\, R_2$ to obtain reduced row-echelon form

$$R = \begin{bmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Since there are two non-zero rows (or equivalently, two leading 1's) the row space and column space are both two dimensional, so rank (A) = 2. To find the nullity of A, we have to find the dimensions of the solution space of the linear system AX = 0. This system can be solved by reducing the coefficient matrix (it is A here) to reduced row-echelon form R. The corresponding system of linear equations will be,

$$x_1 + x_3 + 2x_4 + x_5 = 0$$

$$x_2 + x_3 + x_4 + 2x_5 = 0$$

or, on solving in terms free variable,

$$x_1 = -x_3 - 2x_4 - x_5$$

$$x_2 = -x_3 - x_4 - 2x_5$$

It follows that the general solution of the system is,

$$x_1 = -r - 2s - t$$

$$x_2 = -r - s - 2t$$

$$x_3 = t$$

$$x_4 = s$$

$$x_5 = t$$

Or equivalently,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = r\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \; r, s, t \in \mathbb{R}$$

The three vectors (–1, –1, 1, 0, 0), (–2, –1, 0, 1, 0) and (–1, –2, 0, 0, 1) form a basis for the solution space, so nullity (A) = 3.

Note : In the above example, the given matrix A is 4×5, so that the number of columns of A is n = 5.

We observe that,

rank (A) + nullity (A) = 2 + 3 = 5, the number of columns of A.

Theorem 22 : Dimension Theorem for Matrices

If A is a matrix with n columns, then,

rank (A) + nullity (A) = n

Remark : From the theorem 20 and the foregoing we observe that,

rank (A) = number of leading variables that occur in solving AX = 0.

nullity (A) = number of parameters in the general solution of AX = 0.

Example 3.64 : The matrix

$$A = \begin{bmatrix} 1 & 4 & 5 & 6 & 9 \\ 3 & -2 & 1 & 4 & -1 \\ -1 & 0 & -1 & -2 & -1 \\ 2 & 3 & 5 & 7 & 8 \end{bmatrix}$$

has 5 columns, so

rank A + nullity (A) = 5

This is consistent with example (1), where we observe rank (A) = 2 and nullity (A) = 3.

Example 3.65 : Find nullity of the matrix A, if A is 4×6 and rank (A) = 3.

Solution : From dimension theorem for matrices

$$\begin{aligned} \text{nullity (A)} &= 6 - \text{rank (A)} \\ &= 6 - 3 = 3. \end{aligned}$$

Result (3) : A system of linear equations AX = B is consistent if and only if B is in the column space of A.

Proof : From chapter (1), if $C_1, C_2, \ldots\ldots, C_n$ are columns of A, AX is a linear combination of the column vectors of A. Hence AX = B will hold if and only if B is linear combination of the column vectors of A. Hence AX = B is consistent if and only if B is in the column space of A.

Theorem 23 : The system of linear equations AX = B has a solution if and only if rank A = rank [A | B]; that is, if and only if the ranks of the coefficient matrix and augmented matrix are equal.

Proof : Suppose that $A = [a_{ij}]$ is $m \times n$ matrix, then the given system of linear equations may be written as

$$x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \ldots (i)$$

AX = B has a solution, then there exist values of $x_1, x_2, \ldots, x_n$ that satisfy equation (i). Thus, showing that B is a linear combination of the column of A and hence belongs to the column space of A. Hence, the dimension of the column space of A and [A | B] are equal, so rank A = rank [A|B].

Conversely, suppose that rank A = rank [A|B]. Then B is in the column space of A, which means that we can find values $x_1, x_2, \ldots, x_n$ that satisfy the equation (i). Hence, AX = B has a solution.

Remark :

(1) The above theorem implies that AX = B is in consistent if and only if b is not in the column space of A.

(2) The theorem (23) is infact the consequence of the theorem 24.

Below we give the consolidated list of equivalent statements we proved in the present chapter (for ready reference) for the $n \times n$ matrix A.

1. A is non-singular.
2. X = 0 is the only solution to AX = 0.
3. A is row equivalent to I_n.
4. The linear system of equations AX = B has a unique solution for every $n \times 1$ matrix B.
5. det $A \neq 0$.
6. A has rank n.
7. A has nullity zero.
8. The rows of A form a linearly independent set of n vector in $\mathbb{R}^n$.
9. The columns of A form a linearly independent set of n vectors in $\mathbb{R}^n$.

Example 3.66 : Determine whether $B = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$ is in column space of $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$, if so, express B as linear combination of the columns of A.

Solution : We solve AX = B by Gauss - elimination method. For this consider the augmented matrix (A : B) and reduce it to row-echelon form.

$$[A : B] = \left[\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 3 & 2 \end{array}\right]$$

Apply $R_2 + (-1)\, R_1$ and $R_3 + (-2)R_1$ to obtain

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & -1 & -1 & 1 \\ 0 & -1 & -1 & 4 \end{array}\right]$$

Apply $(-1)\, R_2$ and then $R_2 + (+1)\, R_2$, to obtain,

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 3 \end{array}\right]$$

Since the last row involves all zeros except in last column which is 3, not zero, hence, the system AX = B is inconsistent. That is, there is no solution for AX = B in present case. Therefore, by result (3), B is not linear combination of column vector of A. Hence, B is not in column space of A.

Example 3.67 : Determine whether $B = \begin{bmatrix} 5 \\ 1 \\ -1 \end{bmatrix}$ is in the column space of A, and if so, express B as a linear combination of the column vectors of A, where,

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 9 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Solution : Let AX = B be the linear system.

$$\begin{bmatrix} 1 & -1 & 1 \\ 9 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -1 \end{bmatrix}$$

We solve this system by Gauss-elimination method. For this consider the augmented matrix

$$[A \mid B] = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 9 & 3 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{array}\right]$$

Apply $R_2 + (-9)\, R_1$ and $R_3 + (-1)\, R_1$ to obtain,

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 12 & -8 & -44 \\ 0 & 2 & 0 & -6 \end{array}\right]$$

Apply $\left(\frac{1}{12}\right) R_2$ and then apply $R_3 + (-2)\, R_2$ to obtain,

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 0 & 1 & -\frac{2}{3} & -\frac{11}{3} \\ 0 & 0 & \frac{4}{3} & \frac{4}{3} \end{array}\right]$$

The corresponding system of equations of the last matrix is

$x_1 - x_2 + x_3 = 5$, $x_2 - \frac{2}{3} x_3 = -\frac{1}{3}$ and $\frac{4}{3} x_3 = \frac{4}{3}$.

Solving by back substitution we get, $x_1 = 1$, $x_2 = -3$, and $x_3 = 1$. Thus, the system $AX = B$ has a solution, that is, it is consistent, hence by result (3), B is in the column space of A. Moreover from result (1) we have,

$$\begin{bmatrix} 1 \\ 9 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -1 \end{bmatrix}$$

Example 3.68 : Suppose that $x_1 = -1$, $x_2 = 2$, $x_3 = 4$, $x_4 = -3$ is a solution of a non-homogeneous linear system $AX = b$ and that the solution set of the homogeneous system $AX = 0$ is given by the formulae,

$x_1 = -3r + 4s$, $x_2 = r - s$, $x_3 = r$, $x_4 = s$.

(a) Find the vector form of the general solution of $AX = 0$.

(b) Find the vector form of the general solution of $AX = B$.

Solution : (a) The vector form of the general solution of $AX = 0$ is,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3r + 4s \\ r - s \\ r \\ s \end{bmatrix} = r \begin{bmatrix} -3 \\ 1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 4 \\ -1 \\ 0 \\ 1 \end{bmatrix} \quad \ldots(*)$$

(b) The general solution of AX = B is given by

$$X = X_0 + k_1\vec{u_1} + k_2\vec{u_2},$$

where, X_0 is a particular solution of AX = B. In present case it is given by,

$$X_0 = \begin{bmatrix} -1 \\ 2 \\ 4 \\ -3 \end{bmatrix}$$

and $k_1\vec{u_1} + k_2\vec{u_2}$ is the general solution of AX = 0, given by (∗).

Therefore, the general solution of AX = B is given by,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 4 \\ -3 \end{bmatrix} + k_1\begin{bmatrix} -3 \\ 1 \\ 1 \\ 0 \end{bmatrix} + k_2\begin{bmatrix} 4 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 - 3k_1 + 4k_2 \\ 2 + k_1 - k_2 \\ 4 + k_1 \\ -3 + k_2 \end{bmatrix}$$

Exercise (3.6)

1. Find a basis for the null space of A.

(a) $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & -1 & -3 \\ -1 & 2 & -3 \\ 1 & 1 & 4 \end{bmatrix}$

(c) $A = \begin{bmatrix} 0 & 1 & 3 & -2 \\ 2 & 1 & -4 & 3 \\ 2 & 3 & 2 & -1 \\ -4 & -3 & 5 & -4 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 1 & 4 & 2 & 0 \\ 0 & -2 & -2 & -1 \\ 2 & -4 & 1 & 1 \\ 1 & -2 & -1 & 1 \end{bmatrix}$

(e) $A = \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ -1 & -1 & 2 & -3 & 1 \\ 1 & 1 & -2 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$

2. Find a basis for the row space of A by reducing the matrix to row-echelon form.

(a) $A = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 2 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$

3. For the matrices in exercise (2), find a basis for the column space of A.
4. For the matrices in Exercise (2), find the basis for the row space of A, consisting entirely of row vectors of A.
5. Find a basis for the subspace of $\mathbb{R}^4$ spanned by the given vectors.
   (a) (1, 1, 0, –1), (1, 2, 3, 0), (2, 3, 3, –1)
   (b) (1, 2, 2, –2), (2, 3, 2, –3), (1, 3, 4, –3)
   (c) (1, 1, 0, –1), (1, 2, 3, 0), (2, 3, 3, –1), (1, 2, 2, –2), (2, 3, 2, –3), (1, 3, 4, –3)
6. Find the subset of the vectors that forms a basis for the space spanned by the vectors; then express each vector that is not in the basis as a linear combination of the basis vectors.

   (a) $\overrightarrow{u_1} = (1, 0, 1, 1)$, $\overrightarrow{u_2} = (-3, 3, 7, 1)$, $\overrightarrow{u_3} = (-1, 3, 9, 3)$, $\overrightarrow{u_4} = (-5, 3, 5, -1)$.

   (b) $\overrightarrow{u_1} = (1, -1, 5, 2)$, $\overrightarrow{u_2} = (-2, 3, 1, 0)$, $\overrightarrow{u_3} = (4, -5, 9, 4)$,

   $\overrightarrow{u_4} = (0, 4, 2, -3)$, $\overrightarrow{u_5} = (-7, 18, 2, -8)$
7. Find the rank and nullity of the matrix, then verify the values obtained to satisfy the formula for the dimension theorem.

   (a) $A = \begin{bmatrix} 1 & -1 & 3 \\ 5 & -4 & -4 \\ 7 & -6 & 2 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$

   (c) $A = \begin{bmatrix} 1 & 4 & 5 & 2 \\ 2 & 1 & 3 & 0 \\ -1 & 3 & 2 & 2 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & -3 & 2 & 2 & 1 \\ 0 & 3 & 6 & 0 & -3 \\ 2 & -3 & -2 & 4 & 4 \\ 3 & -6 & 0 & 6 & 5 \\ -2 & 9 & 2 & -4 & -5 \end{bmatrix}$

   (e) $A = \begin{bmatrix} -1 & 2 & 0 & 4 & 5 & -3 \\ 3 & -7 & 2 & 0 & 1 & 4 \\ 2 & -5 & 2 & 4 & 6 & 1 \\ 4 & -9 & 2 & -4 & -4 & 7 \end{bmatrix}$

8. Determine whether B is in the column space of A, and if so, express B as linear combination of the column vectors of A.

(a) $A = \begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix}, B = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix}, B = \begin{bmatrix} 8 \\ 7 \end{bmatrix}$

(c) $A = \begin{bmatrix} 4 & 0 & -1 \\ 3 & 6 & 2 \\ 0 & -1 & 4 \end{bmatrix}, B = \begin{bmatrix} -13 \\ 22 \\ 17 \end{bmatrix}$ (d) $A = \begin{bmatrix} -3 & 6 & 2 \\ 5 & -4 & 0 \\ 2 & 3 & -1 \\ 1 & 8 & 3 \end{bmatrix}, B = \begin{bmatrix} 25 \\ -13 \\ -1 \\ 30 \end{bmatrix}$.

Answers (3.5)

1. (a) Null space = {(0, 0, 0)} (b) Null space = {(0, 0, 0)}.
   (c) {(7, –6, 2, 0), (–5, 4, 0, 2)} (d) Null space = {(0, 0, 0, 0)}.
   (e) {(–1, 1, 0, 0, 0), (–1, 0, –1, 0, 1)}.
2. (a) {(1, –1, 3), (0, 1, –19)} (b) $\left\{(1, 0, -\frac{1}{2})\right\}$
   (c) {(1, 4, 5, 2), (0, 1, 1, 4)}. (d) {(1, –3, 2, 2, 1), (0, 1, 2, 0, –1), (0, 0, 1, 0, –5/12)}
3. (a) $\left\{\begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ -6 \end{bmatrix}\right\}$ (b) $\left\{\begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}\right\}$

   (c) $\left\{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}\right\}$ (d) $\left\{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ -3 \\ -6 \\ 9 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ -2 \\ 0 \\ 2 \end{bmatrix}\right\}$
4. (a) {(1, –1, 3), (5, –4, –4)} (b) {(2, 0, –1)}
   (c) {(1, 4, 5, 2), (2, 1, 3, 0)} (d) {(1, –3, 2, 2, 1), (0, 3, 6, 0 – 3), (2, –3, –2, 4, 4)}.
5. (a) {(1, 1, 0, –1), (0, 1, 3, 1)} (b) {(1, 2, 2, –2), (0, –1, –2, 1)}
   (c) {(1, 1, 0, –1), (0, 1, 3, 1), (0, 0, –1, –2)}.
6. (a) $\{\vec{u_1}, \vec{u_2}\}$; $\vec{u_3} = 2\vec{u_1} + \vec{u_2}$, $\vec{u_4} = -2\vec{u_1} + \vec{u_2}$
   (b) $\{\vec{u_1}, \vec{u_2}, \vec{u_4}\}$; $\vec{u_3} = 2\vec{u_1} - \vec{u_2}$, $\vec{u_5} = -\vec{u_1} + 3\vec{u_2} + 2\vec{u_4}$.
7. (a) rank (A) = 2, nullity (A) = 1; (b) rank (A) = 1, nullity (A) = 2.
   (c) rank (A) = 2, nullity (A) = 2; (d) rank (A) = 3, nullity (A) = 2.
   (e) rank (A) = 2, nullity (A) = 4.

8. (a) $\begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix} = 2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix}$ (b) $\begin{bmatrix} 8 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

(c) $\begin{bmatrix} -13 \\ 22 \\ 17 \end{bmatrix} = (-2)\begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix} + 5\begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix}$.

(d) $\begin{bmatrix} 25 \\ -13 \\ -1 \\ 30 \end{bmatrix} = (-1)\begin{bmatrix} -3 \\ 5 \\ 2 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 6 \\ -4 \\ 3 \\ 8 \end{bmatrix} + 5\begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix}$.

Important Points

- Vector space axioms and consequences of axioms.
- Subspace of a vector space, necessary and sufficient conditions for a subset to be a subspace.
- Linear combination and spanning of vectors.
- Linear span of vectors and set of vectors which span the standard vector spaces.
- Result 1 is known as minus theorem.
- Testing the given set of vectors for linearly dependence or independence.
- Wronskian to test linear dependence of vectors in a function space.
- Basis and dimension of a vector space.
- Finite and infinite dimensional vector spaces.
- Bases of some standard vector spaces.
- Testing the set of vectors for basis of a vector space.
- Row space and column space, of a matrix.
- Rank and nullity of a matrix and the rank-nullity theorem.
- Algorithm for finding rank and nullity of a matrix.

Miscellaneous Exercise

(A) Theory Questions

1. In a vector space $\vec{u}$ a vector in V, and k a scalar, if $k\vec{u} = \vec{0}$, then prove that either $k = 0$ or $\vec{u} = \vec{0}$.
2. Prove that a line passing through the origin in $\mathbb{R}^3$ is a vector space under the standard operations of $\mathbb{R}^3$.
3. State and prove the necessary and sufficient conditions for a subset of a vector space to be a subspace.
4. If $AX = 0$ is a homogeneous linear system of m equations in n unknowns, then prove that the set of solution vectors is a subspace of $\mathbb{R}^n$.

5. If $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ are vectors in a vector space V, then prove that
   (a) The set W of all linear combinations of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ is a subspace of V
   (b) W is the smallest subspace of V containing $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$.
6. Prove that intersection of two subspaces of a vector space V is also a subspace of V.
7. If S is a set with two or more vectors in a vector space, then prove that :
   (a) S is linearly dependent if and only if at least one of the vectors in S is expressible as a linear combination of the other vectors in S.
   (b) S is linearly independent if and only if no vector in S is expressible as a linear combination of the other vectors in S.
8. In $\mathbb{R}^n$ prove that any set S with more than n vectors is linearly dependent.
9. If $B = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ is a basis for a vector space V, then prove that every vector $\vec{v}$ in V is uniquely expressed as a linear combination of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$.
10. Let V be a finite dimensional vector space and let $B = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be any basis.
    (a) If a set S of V has more than n vectors, then prove that S is linearly dependent.
    (b) If a set S of V has fewer than n vectors, then prove that it does not span V.
11. If V is a finite dimensional vector space, then prove that any two bases for V have the same number of vector.
12. If S is a linearly independent set of vectors in a vector space V and if $\vec{v}$ is a vector in V, which is not in L(S), then prove that $S \cup \{\vec{v}\}$ is still linearly dependent.
13. If We is a subspace of a finite dimensional vector space V, then dim (W) ≤ dim (V), moreover, dim (W) = dim (W) if and only if W = V.

(B) Numerical Problems

1. Let V be the set of 2×2 matrices of the form $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}$ with addition defined as $\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix} + \begin{bmatrix} c & 1 \\ 1 & d \end{bmatrix} = \begin{bmatrix} a+c & 1 \\ 1 & b+d \end{bmatrix}$ and scalar multiplication is defined by $k\begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix} = \begin{bmatrix} ka & 1 \\ 1 & kb \end{bmatrix}$, then show that V is a vector space.
2. Find the solution space of the following linear systems :

   (a) $\begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

3. Find the basis and dimensions of the solution spaces in problem 2.
4. Let $\vec{u} = (1, 2, -1)$ and $\vec{v} = (6, 4, 2)$ in $\mathbf{R}^3$. Show that $\vec{w} = (9, 2, 7)$ is linear combination of $\vec{u}$ and $\vec{v}$ and $\vec{w'} = (4, -1, 8)$ is not linear combination of $\vec{u}$ and $\vec{v}$.
5. Show that the vectors $\vec{v_1} = (1, 1, 2)$, $\vec{v_2} = (1, 0, 1)$ and $\vec{v_3} = (2, 1, 3)$ do not span $\mathbf{R}^3$.
6. Determine whether the vectors $\vec{v_1} = (1, -2, 3)$, $\vec{v_2} = (5, 6, -1)$ and $\vec{v_3} = (3, 2, 1)$ form a linearly dependent set or linearly independent set.
7. Use Wronskian to show that the set of vectors $P_1 = 2$, $P_2 = 2 + x$, $P_3 = 2 + x^2$, $P_4 = 2 + x^3$ is linearly independent in $P_3(x)$. Do these vectors form a basis for $P_3(x)$.
8. Let $B = \{(1, 2, 1), (2, 9, 0), (3, 3, 4)\}$.
   (a) Show that B is a basis for $\mathbb{R}^3$.
   (b) Find the co-ordinate vector of $\vec{v} = (5, -1, 9)$ with respect to B.
   (c) Find the vector $\vec{v}$ in $\mathbf{R}^3$ whose co-ordinate vector with respect to B is $(\vec{v})_B = (-1, 3, 2)$.
9. Find the basis and dimension of the solution space of the linear system.

$$\begin{aligned} 2x_1 + 2x_3 - x_3 + x_5 &= 0 \\ -x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 - x_5 &= 0 \\ x_3 + x_4 + x_5 &= 0 \end{aligned}$$

10. Find the basis for the row and column spaces of

$$A = \begin{bmatrix} 1 & -3 & 4 & -2 & 5 & 4 \\ 2 & -6 & 9 & -1 & 8 & 2 \\ 2 & -6 & 9 & -1 & 9 & 7 \\ -1 & 3 & -4 & 2 & -5 & -4 \end{bmatrix}$$

11. Find the basis for the space spanned by the vectors

$\vec{v_1} = (1, -2, 0, 0, 3)$, $\vec{v_2} = (2, -5, -3, -2, 6)$,
$\vec{v_3} = (0, 5, 15, 10, 0)$, $\vec{v_4} = (2, 6, 18, 3, 6)$.

12. Find a subset of the vectors :

$\vec{v_1} = (1, -2, 0, 3)$, $\vec{v_2} = (2, -5, -3, 6)$, $\vec{v_3} = (0, 1, 3, 0)$,

$\vec{v_4} = (2, -1, 4, -7)$, $\vec{v_5} = (5, -8, 1, 2)$.

that form a basis for the space spanned by these vectors.

Answers

2. (a) The solution space is the set of vectors : $x = 2s - 3t$, $y = s$, $z = t$, $s, t \in \mathbb{R}$.
   (b) The solution space is $\{(0, 0, 0)\}$.
   (c) The solution space is the set of vectors : $x = s$, $y = t$, $z = r$, $s, t, r \in \mathbb{R}$.
3. (a) The basis of the solution space is $\{(2, 1, 1), (-3, 0, 1)\}$ and dimension = 2.
   (b) The solution space is trivial subspace of $\mathbb{R}^3$.
   (c) The basis of the solution space is $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.
4. $\vec{W} = -3\vec{u} + 2\vec{v}$.
5. Determinant of the matrix whose columns are $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$ is zero.
6. Linearly dependent for $-1\vec{v_1} - 1 \cdot \vec{v_2} + 2\vec{v_2} = \vec{0}$.
7. Yes, they form a basis for $P_3(x)$.
8. Yes, they form a basis for $P_3(x)$.
9. (b) $(\vec{v})_B = (1, -1, 2)$, (c) $\vec{v} = (11, 31, 7)$.
10. Basis is $\{(-1, 1, 0, 0, 0), (-1, 0, -1, 0, 1)\}$ and dimension is 2.
11. Basis for row space of A is $\{(1, -3, 4, -2, 5, 4), (0, 0, 1, 3, -2, -6), (0, 0, 0, 0, 1, 5)$ dimension of the row space is 3.
    The basis for column space of A is $\{(1, 2, 2, -1), (4, 9, 9, -4), (5, 8, 9, -5)$ and dimension of the column space is also 3.
12. The non-zero row vectors is row-echelon form of A are $\vec{w_1} = (1, -2, 0, 0, 3)$, $\vec{w_2} = (0, 1, 3, 2, 0)$, $\vec{w_3} = (0, 0, 1, 1, 0)$. These form a basis for the row space and consequently form a basis for the subspace of $\mathbb{R}^5$ spanned by $\vec{v_1}$, $\vec{v_2}$, $\vec{v_3}$ and $\vec{v_n}$.
13. Consider a matrix whose column vectors are the vectors $\vec{v_1}$, $\vec{v_2}$, $\vec{v_3}$, $\vec{v_4}$ and $\vec{v_5}$ and reduce it to reduced row-echelon form.

    Leading 1's occur in columns 1, 2 and 4, so the column vectors $\vec{w_1}$, $\vec{w_2}$ and $\vec{w_4}$ form a basis for the column space of A, where $\vec{w_1} = (1, 0, 0, 0)$, $\vec{w_2} = (0, 1, 0, 0)$ and $\vec{w_4} = (0, 0, 1, 0)$ and consequently, $(v_1, v_2, v_4\}$ is a basis for the space spanned by the given system.

❐❐❐

Chapter **4**...

Eigenvalues and Eigenvectors

Contents

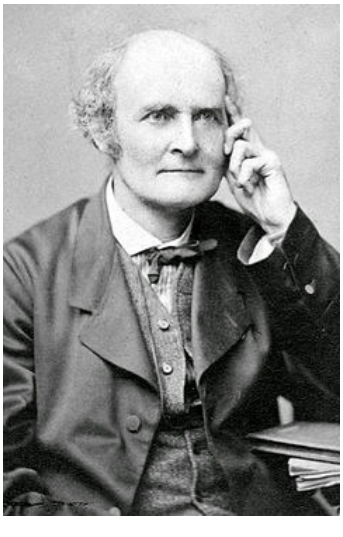
4.1 Introduction
4.2 Characteristics Polynomials of Degree 2 and 3
4.3 Cayley Hamilton Theorem
4.4 The Construction of Orthogonal Matrices
4.5 Diagonalization of Matrices
4.6 Minimal Polynomial and Minimal Equation of a Matrix
4.7 Derogatory and Non-derogatory Matrices

Objectives

At the end of this chapter.

(1) Student should understand the concept of diagonalization and triangularization of a matrix.

(2) Student should understand the concept of characteristics polynomial and minimal polynomial.



Arthur Cayley F.R.S. (/ˈkeɪli/; 16 August 1821 – 26 January 1895) was a British mathematician. He helped found the modern British school of pure mathematics.

As a child, Cayley enjoyed solving complex maths problems for amusement. He entered Trinity College, Cambridge, where he excelled in Greek, French, German, and Italian, as well as mathematics. He worked as a lawyer for 14 years.

He postulated the Cayley-Hamilton theorem that every square matrix is a root of its own characteristic polynomial, and verified it for matrices of order 2 and 3. He was the first to define the concept of a group in the modern way as a set with a binary operation satisfying certain laws. Formerly, when mathematicians spoke of "groups", they had meant permutation groups.

Arthur Cayley

William Rowan Hamilton

Sir William Rowan Hamilton PRIA FRSE (4 August 1805 – 2 September 1865) was an Irish physicist, astronomer, and mathematician, who made important contributions to classical mechanics, optics, and algebra. His studies of mechanical and optical systems led him to discover new mathematical concepts and techniques. His best known contribution to mathematical physics is the reformulation of Newtonian mechanics, now called Hamiltonian mechanics. This work has proven central to the modern study of classical field theories such as electromagnetism, and to the development of quantum mechanics. In pure mathematics, he is best known as the inventor of quaternions.

4.1 Introduction

If A is an $n \times n$ matrix and X is a vector in $\mathbb{R}^n$ (X considered as $n \times 1$ column matrix), we are going to study the properties of non-zero vector X, where X and AX are scalar multiples of one another. Such vectors X arise naturally in the study of vibrations, electrical systems, genetics, chemical reactions, quantum mechanics, economics, and geometry. In this section, we shall see how to find these vectors.

EIGENVALUES AND EIGENVECTORS OF MATRICES

Definition : If A is an $n \times n$ matrix, then a *non-zero* vector X in $\mathbb{R}^n$ is called *eigenvector of A* if

$$AX = \lambda X$$

for some scalar λ. The scalar λ is called an *eigenvalue of A*, and X is said to be an eigenvector of A corresponding to eigenvalue λ.

Remark : Eigenvalues are also called *proper values* or *characteristic values* (roots), latent and eivenvectors are called characteristic vector or latent vectors.

SOLVED EXAMPLE

Example 4.1 : The vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$.

Solution : Since we have,

$$AX = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

That is, $AX = 3X$

Therefore, $X = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of A corresponding to eigenvalue $\lambda = 3$.

Theorem 1 : If A is an $n \times n$ matrix and λ is a real number, then λ is an eigenvalue of A if and only if $\det(\lambda I - A) = 0$.

Proof : If λ is an eigenvalue of A, then there exists a non-zero X in $\mathbb{R}^n$ such that,

$$AX = \lambda X \qquad \ldots\text{(By definition)}$$

We write this equation as,

$$AX = \lambda IX, \qquad \ldots \text{ where I is an } n \times n \text{ unit matrix}$$

or equivalently,

$$-AX + \lambda IX = 0$$

or
$$(\lambda I - A)X = 0 \qquad \ldots(1)$$

For λ to be an eigenvalue of A, there must be a non-zero solution of equation (1). We know that, if A is an $n \times n$ matrix and X is in $\mathbb{R}^n$ then $AX = 0$, has non-trivial solution if and only if $\det A = 0$. Therefore, the equation (1) will have non-zero solution if and only if,

$$\det(\lambda I - A) = 0$$

Conversely, if $\det(\lambda I - A) = 0$ then by the same result there will be a non-zero solution for the equation,

$$(\lambda I - A)X = 0$$

That is, there will be a non-zero X in $\mathbb{R}^n$ such that $AX = \lambda X$, which shows that λ is an eigenvalue of A.

Definition : If A is an $n \times n$ matrix and λ is a real number, then the equation,

$$\Delta(\lambda) = \det(\lambda I - A) = 0$$

is called the *characteristic equation of A.* When expanded, the determinant $\det(\lambda I - A)$ is a polynomial in λ called the *characteristic polynomial of A.*

If A is an $n \times n$ matrix, then the characteristic polynomial of A has degree n and the coefficient of λ^n is 1. Thus, the characteristic polynomial of an $n \times n$ matrix has the form,

$$\Delta(\lambda) = \det(\lambda I - A) = \lambda^n + C_1 \lambda^{n-1} + \ldots\ldots + C_n.$$

4.2 Characteristic Polynomial of Degree 2 and 3

There are simple formulae for the characteristics polynomial of matrices of order 2 and 3.

(a) Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ then

$$\Delta(t) = t^2 - (a_{11} + a_{22})\, t + \det(A)$$
$$= t^2 - \text{tr}(A)\, t + \det(A)$$

Here tr (A) denotes the trace of A, that is the sum of diagonal elements of A.

(b) Suppose $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ then

$$\Delta(t) = t^3 - \text{tr}(A)\, t^2 + (A_{11} + A_{12} + A_{33})\, t - \det(A)$$

(Here A_{11}, A_{22}, A_{33} denote respectively the cofactors of a_{11}, a_{22}, a_{33})

Example : Find the characteristic polynomials of each of the following matrices :

(a) $A = \begin{bmatrix} 7 & -1 \\ 2 & 6 \end{bmatrix}$ $B = \begin{bmatrix} 5 & 2 \\ 4 & -4 \end{bmatrix}$

(a) We have, $\text{tr}(A) = 7 + 6 = 13$ and $|A| = 42 + 2 \Rightarrow 44$

Hence, $\Delta(t) = t^2 - 13t + 44$

(b) We have, $\text{tr}(B) = 5 - 4 = 1$ and $|B| = -20 - 8 = -28$

Hence, $\Delta(t) = t^2 - t - 28$

For example : Find the characteristics polynomials of $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}$.

We have, $\text{tr}(A) = 1 + 3 + 9 = 13$

The cofactors of the diagonal elements are as follows :

$$A_{11} = \begin{vmatrix} 3 & 2 \\ 3 & 9 \end{vmatrix} = 21, \quad A_{22} = \begin{vmatrix} 1 & 2 \\ 1 & 9 \end{vmatrix} = 7, \quad A_{33} = \begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} = 3.$$

Thus, $A_{11} + A_{22} + A_{33} = 31$, also $|A| = 17$.

$\therefore$ The characteristic polynomial

$$\Delta(t) = t^3 - 13t^2 + 31t - 17$$

Note : The coefficients of the characteristic polynomial $\Delta(t)$ of square matrix A having order 3, with alternating signs, as follows :

$$s_1 = \text{tr}(a)$$
$$s_2 = A_{11} + A_{22} + A_{33}$$
$$s_3 = \det(A)$$

Note that : Each s_k is the sum of all principal minors of A of order k.

Theorem 2 : Let A be n-square matrix then it's characteristic polynomial is

$$\Delta(t) = t^n - s_1 t^{n-1} + s_2 t^{n-2} + \ldots + (-1)^n s_n$$

where s_k is the sum of the principle minors of order k.

SOLVED EXAMPLES

Example 4.2 : Find the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 7 \\ 1 & -2 \end{bmatrix}$$

Solution : Since,

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 7 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} \lambda - 2 & -7 \\ -1 & \lambda + 2 \end{bmatrix}$$

the characteristic polynomial of A is,

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -7 \\ -1 & \lambda + 2 \end{bmatrix} = \lambda^2 - 11.$$

and the characteristic equation of A is,

$$\lambda^2 - 11 = 0$$

The solutions of this equation are $\lambda = +\sqrt{11}$ and $\lambda = -\sqrt{11}$; these are the eigenvalues of A.

Example 4.3 : Find the eigenvalues of the matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

Solution : We have,

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -2 & \lambda - 3 & -4 \\ 1 & 1 & \lambda + 2 \end{bmatrix}$$

Therefore, the characteristic polynomial of A is,

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -2 & \lambda - 3 & -4 \\ 1 & 1 & \lambda + 2 \end{bmatrix}$$

$$= (\lambda - 2)[(\lambda - 3)(\lambda + 2) + 4] + 1[-2(\lambda + 2) + 4] - 1[-2 - (\lambda - 3)]$$

$$= (\lambda - 2)(\lambda^2 - \lambda - 2) - 2\lambda + 2 + \lambda - 3$$

$$= \lambda^3 - 3\lambda^2 - \lambda + 3.$$

Thus, $\det(\lambda I - A) = \lambda^3 - 3\lambda^2 - \lambda + 3$.

The characteristic equation of A is,

$$\lambda^3 - 3\lambda^2 - \lambda + 3 = 0$$

or $(\lambda - 1)(\lambda + 1)(\lambda - 3) = 0$

The solutions of this equation are $\lambda = 1$, $\lambda = -1$, and $\lambda = 3$. Therefore, $\lambda = 1$, $\lambda = -1$ and $\lambda = 3$ are the eigenvalues of A.

From the definition of eigenvalues and eigenvectors of a matrix and theorem 1 (above), the results of the following theorem are obvious and can be proved easily, hence, it is left as an exercise.

Theorem 3 : If A is an $n \times n$ matrix and λ is a real number, then the following are equivalent :

(a) λ is an eigenvalue of A.

(b) The system of equations $(\lambda I - A)X = 0$ has non-trivial solutions.

(c) There is a non-zero vector X in $\mathbb{R}^n$ such that $AX = \lambda X$.

(d) λ is a solution of the characteristic equation $\det(\lambda I - A) = 0$.

Eigenspace of matrix :

Definition : Let A be an $n \times n$ matrix and λ be the eigenvalue of A. The set of all vectors X in $\mathbb{R}^n$ which satisfy the identity $AX = \lambda X$ is called the *eigenspace* of A corresponding to λ. This is denoted by $E(\lambda)$.

Remark : The eigenvectors of A corresponding to an eigenvalue λ are the non-zero vectors X that satisfy $AX = \lambda X$. Equivalently, the eigenvectors corresponding to λ are the non-zero vectors in the solution space of $(\lambda I - A)\, X = 0$. Therefore, the eigenspace is the set of all non-zero X that satisfy $(\lambda I - A)\, X = 0$ with trivial solution in addition.

Our main objective in this section is to find eigenvalues and eigenvectors of a given $n \times n$ matrix. We describe in the below step-by-step working method for finding eigenvalues and the corresponding eigenspace.

Let A be an $n \times n$ matrix.

Step 1 : For real number λ, form the matrix $\lambda I - A$.

Step 2 : Evaluate $\det(\lambda I - A)$; that is, the characteristic polynomial of A.

Step 3 : Consider the equation $\det(\lambda I - A) = 0$ (the characteristic equation of A), solve this equation for λ.

Let $\lambda_1, \lambda_2, \ldots\ldots, \lambda_n$ be the eigenvalues of A thus calculated.

Step 4 : For each λ_i, consider the equation $(\lambda_i I - A)\, X = 0$

which is a homogeneous system of n linear equations in n unknowns. Find the solution space of this system which is an eigenspace $E(\lambda_i)$ of A corresponding to the eigenvalue λ_i of A. Repeat this for each $i = 1, 2, \ldots\ldots n$.

Step 5 : From step 4, we can find basis and dimension for each eigenspace $E(\lambda_i)$, for $i = 1, 2, \ldots\ldots, n$.

Before we workout examples, let us prove the following simple result.

Theorem 4 : Let A be an $n \times n$ matrix and λ be an eigenvalue of A. Then the eigenspace $E(\lambda)$ is a subspace of $\mathbb{R}^n$.

Proof : Since $E(\lambda) = \{ X \in \mathbb{R}^n | AX = \lambda X\}$, the zero vector belongs to $E(\lambda)$ as $A0 = \lambda 0 = 0$. In order to prove that $E(\lambda)$ is a subspace of $\mathbb{R}^n$, we need to show that $E(\lambda)$ is closed under addition and scalar multiplication. Let X_1 and X_2 be any two vectors in $E(\lambda)$, so that

$$AX_1 = \lambda X_1$$

and $$AX_2 = \lambda X_2.$$

We have, $$A(X_1 + X_2) = AX_1 + AX_2$$

$$= \lambda X_1 + \lambda X_2 = \lambda (X_1 + X_2)$$

that is, $$A(X_1 + X_2) = \lambda (X_1 + X_2),$$

showing that $X_1 + X_2$ is in $E(\lambda)$.

Also for any scalar k,

$$A(kX) = k(AX) = k(\lambda X) = \lambda(kX)$$

that is $A(kX) = \lambda(kX)$, showing that kX is also in $E(\lambda)$.

Thus, $E(\lambda)$ is a subspace of $\mathbb{R}^n$.

SOLVED EXAMPLES

Example 4.4 : Find : (a) characteristic polynomial, (b) eigenvalues and (c) basis for the eigenspaces of a matrix

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

Solution : We have, for the real number λ,

$$\lambda I - A = \begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix}$$

So the characteristic polynomial of A is

$$\det(\lambda I - A) = \det\begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix}$$

$$= (\lambda - 3)(\lambda + 1)$$

The eigenvalues of A are given by the characteristic equation

$$\det(\lambda I - A) = 0$$

Or equivalently,

$$(\lambda - 3)(\lambda + 1) = 0$$

which shows that the eigenvalues of A are 3 and –1.

To find eigenspace of A corresponding to each eigenvalue of A, we have to find solution space of,

$$(\lambda I - A)X = \begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \ldots(*)$$

(i) For $\lambda = 3$, the system (*) becomes,

$$\begin{bmatrix} 0 & 0 \\ -8 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or equivalently, $-8x_1 + 4x_2 = 0$

or $-2x_1 + x_2 = 0$

The solution of this system is given by

$$x_1 = \frac{1}{2}t,\ x_2 = t,\ t \in \mathbb{R}$$

Then the eigenspace E(λ) of A corresponding to λ = 3, is given by,

$$E(\lambda) = \left\{ \left(\frac{1}{2}t, t\right) \middle| t \in \mathbb{R} \right\}$$

$$= \left\{ t\left(\frac{1}{2}, 1\right) \middle| t \in \mathbb{R} \right\}$$

The basis for E (λ) is $\left\{\left(\frac{1}{2}, 1\right)\right\}$ and dim E(λ) = 1.

(ii) For λ = –1, the system (∗) takes the form, $\begin{bmatrix} -4 & 0 \\ -8 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or equivalently, $-4x_1 = 0$

$-8x_1 = 0$

The solution of this system (remember x_2 is free variable) is

$$x_1 = 0 , x_2 = t, t \in \mathbb{R}$$

$$\therefore \quad E(\lambda = -1) = \{ (0, t) \mid t \in \mathbb{R} \}$$

$$= \{ t(0, 1) \mid t \in \mathbb{R} \}$$

∴ Basis for E (λ) is { (0, 1) } and dim E(λ) = 1.

Example 4.5 : Find the eigenvalues of the matrix,

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

Also find eigenspace corresponding to each eigenvalue of A. Further find basis and dimension for the same.

Solution : In example 4.3, we have calculated the eigenvalues of A, which are given by $\lambda_1 = 1$, $\lambda_2 = -1$, and $\lambda_3 = 3$.

We have,

$$(\lambda I - A)\, X = \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -2 & \lambda - 3 & -4 \\ 1 & 1 & \lambda + 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \ldots(*)$$

(i) For $\lambda = \lambda_1 = 1$ in (*), we have,

$$\begin{bmatrix} -1 & -1 & -1 \\ -2 & -2 & -4 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing $R_2 - (2)\, R_1$, $R_3 + R_1$ and then $(-1)\, R_1$),

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or (by performing $R_3 + R_2$)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of equations corresponding to the last matrix equation is,

$$x_1 + x_2 + x_3 = 0$$
$$-2x_3 = 0$$

The general solution of this system is $x_1 = -t$, $x_2 = t$, $x_3 = 0$, $t \in \mathbb{R}$.

Therefore, for $\lambda = 1$,

$$E(\lambda) = \{ (-t, t, 0) \mid t \in \mathbb{R} \}$$
$$= \{ t(-1, 1, 0) \mid t \in \mathbb{R} \}$$

Thus, the eigenspace of A corresponding to $\lambda = 1$ is the set of all scalar multiples of $(-1, 1, 0)$. So that $\dim E(\lambda) = 1$, and basis for $E(\lambda)$ is $\{(-1, 1, 0)\}$.

(ii) For $\lambda = -1$, in $(*)$, we have,

$$\begin{bmatrix} -3 & -1 & -1 \\ -2 & -4 & -4 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing R_{13}, then perform $R_2 + 2R_1$ and $R_3 + 3R_1$),

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & -2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or (by performing $R_3 + (-1)\, R_2$, then $\left(-\frac{1}{2}\right) R_2$)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of linear equations corresponding to the last matrix equation is,

$$x_1 + x_2 + x_3 = 0$$
$$x_2 + x_3 = 0$$

Solving this system we get,

$$x_1 = 0\ ;\ x_2 = -t,\ x_3 = t\ ,\ t \in \mathbb{R}.$$

Therefore,

$$E(\lambda = -1) = \{ (0, -t, t) \mid t \in \mathbb{R} \}$$
$$= \{ t(0, -1, 1) \mid t \in \mathbb{R} \}$$

Thus, dim $E(\lambda) = 1$ and basis for $E(\lambda)$ is $\{(0, -1, 1)\}$.

(iii) For $\lambda = 3$ in (*), we have,

$$\begin{bmatrix} 1 & -1 & -1 \\ -2 & 0 & -4 \\ 1 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing $R_2 + 2R_1$ and $R_3 + (-1) R_1$)

$$\begin{bmatrix} 1 & -1 & -1 \\ 0 & -2 & -6 \\ 0 & 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or (on performing $R_3 + R_2$ then $\left(-\frac{1}{2}\right) R_2$)

$$\begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of equations corresponding to the last matrix equation is,

$$x_1 - x_2 - x_3 = 0$$
$$x_2 + 3x_3 = 0$$

The solutions of this system are given by $x_1 = -2t$, $x_2 = -3t$, $x_3 = t$, $t \in \mathbb{R}$.

$$\therefore \quad E(\lambda) = \{ (-2t, -3t, t) \mid t \in \mathbb{R} \}$$
$$= \{ t(-2, -3, 1) \mid t \in \mathbb{R} \}$$

Thus, dim $E(\lambda) = 1$, and basis for $E(\lambda)$ is $\{ (-2, -3, 1) \}$.

Example 4.6 : Find eigenvalues, eigenspace, basis and dimension of eigenspace of the matrix

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$$

Solution : The characteristic polynomial of A is,

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & 1 & -1 \\ 0 & \lambda - 3 & 1 \\ -2 & -1 & \lambda - 3 \end{bmatrix}$$
$$= \lambda^3 - 3\lambda^2 + 20\lambda + 16 = (\lambda - 2)(\lambda - 2)(\lambda - 4)$$

So that the eigenvalues of A are 2, 2 and 4.

Now, let us find eigenspace for each λ. For this, we have to solve

$$(\lambda I - A)X = \begin{bmatrix} \lambda-2 & 1 & -1 \\ 0 & \lambda-3 & 1 \\ -2 & -1 & \lambda-3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \ldots(*)$$

(i) For $\lambda = 2$, we have,

$$\begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ -2 & -1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing R_{13}, then $R_3 + R_2$ and $\left(-\frac{1}{2}\right) R_1$)

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of linear equations corresponding to the last equation is,

$$x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = 0$$

$$-x_2 + x_3 = 0$$

Solving these equations, we get,

$$x_1 = -t \; ; \; x_2 = t \; ; \; x_3 = t, \; t \in \mathbb{R}$$

Therefore, $$E(\lambda = 2) = \left\{ (-t, t, t) \middle| t \in \mathbb{R} \right\}$$

$$= \left\{ t(-1, 1, 1) \middle| t \in \mathbb{R} \right\}$$

Thus, dim E (λ) = 1 and basis for E (λ) is $\{(-1, 1, 1)\}$.

(ii) For $\lambda = 4$, in (*), we have,

$$\begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ -2 & -1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing $R_3 + R_1$),

$$\begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of equations corresponding to last matrix equation is,

$$2x_1 + x_2 - x_3 = 0$$

$$x_2 + x_3 = 0$$

Solving we get,

$$x_1 = t \; ; \; x_2 = -t, \; x_3 = t, \; t \in \mathbb{R}.$$

$$\therefore \quad E(\lambda = 4) = \left\{ (t, -t, t) \middle| \; t \in \mathbb{R} \right\}$$

$$= \left\{ t(1, -1, 1) \middle| \; t \in \mathbb{R} \right\}$$

Therefore, dim $(E(\lambda = 4)) = 1$, basis for it is $\{ (1, -1, 1) \}$.

Example 4.7 : Find eigenvalues, eigenspace, dimension and basis for the eigenspace of the matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

Solution : The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -2 & \lambda - 3 & -2 \\ -3 & -3 & \lambda - 4 \end{bmatrix}$$

$$= \lambda^3 - 9\lambda^2 + 15\lambda - 7 = (\lambda - 1)(\lambda - 1)(\lambda - 7)$$

The roots of this polynomial are 1, 1 and 7. Therefore, eigenvalues of A are 1, 1 and 7.

To find eigenspace corresponding to each λ, we have to solve,

$$(\lambda I - A)X = \begin{bmatrix} \lambda - 2 & -1 & -1 \\ -2 & \lambda - 3 & -2 \\ -3 & -3 & \lambda - 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad (*)$$

For each $\lambda = 1, 1, 7$.

(i) For $\lambda = 1$, the equation $(*)$ becomes,

$$\begin{bmatrix} -1 & -1 & -1 \\ -2 & -2 & -2 \\ -3 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing $(-1)\, R_1$, then $R_2 + 2R_1$ and $R_3 + 3R_1$)

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of equations corresponding to the last matrix equation is,

$$x_1 + x_2 + x_3 = 0$$

Solving this we get,

$x_1 = -r - s,\ x_2 = r,\ x_3 = s,\ r, s \in \mathbb{R}$.

Therefore, $E(\lambda) = \{ (-r - s, r, s) \mid r, s \in \mathbb{R} \}$

$= \{ r(-1, 1, 0) + s(-1, 0, 1) \mid r, s \in \mathbb{R} \}$

Here $E(\lambda = 1)$ is the set of vectors which are linear combinations of the vectors which are linearly independent (check !). Therefore, dim $E(\lambda = 1) = 2$, and basis for the same is, $\{(-1, 1, 0), (-1, 0, 1)\}$.

(ii) For $\lambda = 7$, the equation $(*)$ becomes,

$$\begin{bmatrix} 5 & -1 & -1 \\ -2 & 4 & -2 \\ -3 & -3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or equivalently (by performing in sequence R_{13}, then $\left(-\frac{1}{3}\right) R_1$, $R_2 + 2R_1$, $R_3 + (-5) R_1$),

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 6 & -4 \\ 0 & -6 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Or (by performing $R_3 + R_2$)

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 6 & -4 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The system of equations corresponding to last matrix equation is,

$$x_1 + x_2 - x_3 = 0$$
$$6x_2 - 4x_3 = 0$$

Solving we get,

$$x_1 = \frac{1}{3} t\ ;\ x_2 = -\frac{2}{3} t\ ;\ x_3 = t.$$

Therefore,

$$E(\lambda = 7) = \left\{ \left(\frac{1}{3} t, -\frac{2}{3} t, t \right) \middle| \ t \in \mathbb{R} \right\} = \left\{ t \left(\frac{1}{3}, -\frac{2}{3}, 1 \right) \middle| \ t \in \mathbb{R} \right\}$$

Thus, dim $E(\lambda = 7) = 1$, and basis for it is

$$\left\{ \left(\frac{1}{3}, -\frac{2}{3}, 1 \right) \right\}$$

Eigenvalues of triangular matrices :

The determinant of any triangular matrix is the product of the elements on main diagonal.

Example 4.8 : If,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

then the eigenvalues of A are a_{11}, a_{22}, a_{33} and a_{44}.

The characteristic polynomial of A is,

$\det(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44})$

$\therefore$ The eigenvalues of A are the entries on the main diagonal of A.

Theorem 5 : If k is positive integer, λ is an eigenvalue of a matrix A, X is a corresponding eigenvector, then λ^k is an eigenvalue of A^k and X is a corresponding eigenvector.

Proof : We prove this result by mathematical induction. Let the statement be $P_k : A^k \cdot X = \lambda^k X$.

If k = 1, by hypothesis, λ is eigenvalue of A and X is corresponding eigenvector; that is, $AX = \lambda X$, so the result is true for k = 1.

Let us assume the statement is true for k = r ; that is, $A^r X = \lambda^r X$ holds. Then consider,

$$\begin{aligned} A^{r+1}X &= A(A^r X) \\ &= A(\lambda^r X) && \ldots\text{by induction hypothesis} \\ &= \lambda^r (AX) && \because \lambda^r \text{ is a scalar.} \\ &= \lambda^r (\lambda X) \end{aligned}$$

$$\therefore \quad A^{r+1}X = \lambda^{r+1}X$$

Showing that λ^{r+1} is an eigenvalue of A^{r+1} and X is the corresponding eigenvector.

Therefore, the statement is true for k = r + 1. Hence, by mathematical induction, the statement is true for all $k \geq 1$, that is,

$$A^k X = \lambda^k X, \qquad \text{for all integer } k \geq 1.$$

Hence, the proof.

SOLVED EXAMPLES

Example 4.9 : Find the eigenvalues and bases for eigenspaces of A^5, where,

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

Solution : In Example (4.5), we have seen that 1, –1, and 3 are eigenvalues of A. Further (–1, 1, 0) and its non-zero scalar multiples are eigenvectors of A corresponding to the eigenvalue $\lambda = 1$.

(0, –1, 1) and its non-zero scalar multiples are eigenvalues of A corresponding to $\lambda = -1$. (–2, –3, 1) and its non-zero scalar multiples are eigenvectors of A corresponding to $\lambda = 3$.

Then by theorem 4, if λ is eigenvalue of A and X the corresponding eigenvector, then λ^k is eigenvalue of A^k, and the corresponding eigenvector is X, we see that $(1)^5 = 1$, $(-1)^5 = -1$ and $(3)^5 = 243$ are the eigenvalues of A^5 and the corresponding eigenvectors are (–1, 1, 0), (0, –1, 1) and (–2, –3, 1) respectively.

In fact $E(\lambda = 1) = \{ t(-1, 1, 0) \mid t \in \mathbb{R} \}$ with basis $\{ (1, 1, 0) \}$

$E(\lambda = -1) = \{ t(0, -1, 1) \mid t \in \mathbb{R} \}$ with basis $\{(0, -1, 1)\}$

and $E(\lambda = 243) = \{ t(-2, -3, 1) \mid t \in \mathbb{R} \}$, with basis $\{(-2, -3, 1)\}$.

Example 4.10 : Let $A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$ be a matrix. Find the $\Delta(\lambda)$, the characteristic polynomial of A. Further verify that $\Delta(A) = 0$, that is, A satisfies its characteristic polynomial.

Solution : We have the characteristic matrix

$$\lambda I - A = \begin{bmatrix} \lambda - 1 & 3 & -3 \\ -3 & \lambda + 5 & -3 \\ -6 & 6 & \lambda - 4 \end{bmatrix}$$

So the characteristic polynomial $\Delta(\lambda)$ is

$$\Delta(\lambda) = \det(\lambda I - A)$$

$$= \det \begin{bmatrix} \lambda - 1 & 3 & -3 \\ -3 & \lambda + 5 & -3 \\ -6 & 6 & \lambda - 4 \end{bmatrix}$$

$$= (\lambda - 1)[(\lambda + 5)(\lambda - 4) + 18] - 3[-3(\lambda - 4) - 18] - 3[-18 + 6(\lambda + 5)]$$

$$= (\lambda - 1)(\lambda^2 + \lambda - 2) + 9(\lambda - 4 + 6) - 18[-3 + (\lambda + 5)]$$

$$= (\lambda + 2)^2 (\lambda - 4)$$

Thus, $\Delta(\lambda) = (\lambda + 2)^2 (\lambda - 4)$.

Note that an $n \times n$ matrix A satisfies a polynomial

$f(x) = a_0 + a_1 x + \ldots + a_n x^n$ means

$f(A) = a_0 I_n + a_1 A + \ldots + a_n A^n = 0$, where 0 is an $n \times n$ zero matrix.

Therefore, consider

$$\Delta(A) = (A + 2I_3)^2 (A - 4I_3)$$

$$= \left\{ \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} + 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}^2 \left\{ \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} - 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

$$= \begin{bmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{bmatrix}^2 \begin{bmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 18 & -18 & 18 \\ 36 & -36 & 36 \\ 36 & -36 & 36 \end{bmatrix} \begin{bmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The above example (4.10) has the significant generalization known as *Cayley-Hamilton Theorem.*

4.3 Cayley Hamilton Theorem

Statement : Every matrix A is a root of its characteristic polynomials.

OR

Theorem 6 (Cayley-Hamilton Theorem) : Every n × n matrix satisfies its characteristic equation.

OR

Every square matrix is zero of its characteristic polynomial.

OR

Every matrix A is a root of it's characteristic polynomial.

Proof : Let A be an arbitrary n-square matrix and $\Delta(t)$ be it's characteristic polynomial, say $\Delta(t) = |tI - A| = t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$.

Now B(t) denote the classical adjoint of the matrix (tI − A). The element of B(t) are cofactors of the matrix (tI − A) and hence are polynomials in t of degree not exceeding n − 1.

Thus, $B(t) = B_{n-1} t^{n-1} + \ldots + B_1 t + B_0$

where, B_i are n-square matrices over k, which are independent of t.

By the fundamental property of classical adjoint

$(tI - A) B(t) = |tI - A| I$

or $(tI - A)(B_{n-1} t^{n-1} + \ldots + B_1 t + B_0) = (t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0) I$

Removing the parentheses and equating corresponding power of y yields.

$B_{n-1} = I,$

$B_{n-2} - A\,B_{n-1} = a_{n-1}\,I,$

: : :

$B_0 - AB_1 = a_1 I$

$\therefore \quad -AB_0 = a_0 I$

Multiplying the above equations by $A^n, A^{n-1}, \ldots, A, I$ respectively, yields

$A^n\,B_{n-1} = A_n I,$

$A^{n-1}\,B_{n-2} - A^n\,B_{n-1} = a_{n-1}\,A^{n-1}$

: :

$AB_0 - A^2B_1 = a_1 A$

$-AB_0 = a_0 I$

Adding the above matrix equations yields 0 on the left hand side and $\Delta(A)$ on right side, that is

$$0 = A^n + a_{n-1}\,A^{n-1} + \ldots + a_1 A + a_0 I$$

Therefore, $\Delta(A) = 0$, which is the Cayley-Hamilton theorem.

Theorem 7 : If A is n-square matrix, then A and its transpose A^t have the same characteristic polynomial.

Proof : We have,

$$(\lambda I - A)^t = (\lambda I)^t - A^t = \lambda I - A^t$$

We know that a matrix and its transpose have the same determinant. Hence,

$$\det (\lambda I - A)^t = \det (\lambda I - A)$$

But $(\lambda I - A)^t = \lambda I - A^t$ gives

$$\det (\lambda I - A) = \det (\lambda I - A^t)$$

Showing that the characteristic polynomial of A and A^t are same.

Theorem 8 : If A and B are n-square matrices, then AB and BA have the same non-zero eigenvalues.

Proof : Suppose λ is a non-zero eigenvalue of AB. Then there exists a non-zero vector $\overrightarrow{v}$ such that $AB\overrightarrow{v} = \lambda\overrightarrow{v}$. Put $\overrightarrow{w} = B\overrightarrow{v}$. Since $\lambda \neq 0$ and $\overrightarrow{v} \neq \overrightarrow{0}$,

$$A\overrightarrow{w} = A(B\overrightarrow{v}) = (AB)\overrightarrow{v} = \lambda\overrightarrow{v} \neq \overrightarrow{0}$$

and hence $\overrightarrow{w} \neq \overrightarrow{0}$.

Now, $BA\overrightarrow{w} = BA(B\overrightarrow{v}) = (BAB)\overrightarrow{v} = B(AB\overrightarrow{v}) = B\lambda\overrightarrow{v} = \lambda B\overrightarrow{v} = \lambda\overrightarrow{w}$.

Thus, $BA\vec{w} = \lambda\vec{w}$, showing that $\vec{w}$ is an eigenvector of BA corresponding to the eigenvalue λ. That is, λ is an eigenvalue of BA. Similarly, we can show that, any non-zero eigenvalue of BA is also an eigenvalue of AB.

Thus, AB and BA have the same non-zero eigenvalues.

Theorem 9 : A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

Proof : Let A be an $n \times n$ matrix and let $f(x) = x^n + c_1 x^{n-1} + \ldots + c_{n-1} x + c_n$ be the characteristics polynomial of A. We observe that $\lambda = 0$ satisfies f(x), that is $\lambda = 0$ is eigenvalue of A if and only if $f(\lambda = 0) = 0$, which shows that $\lambda = 0$ is eigenvalue of A if and only if $c_n = 0$. Thus, then it suffices to prove that A is invertible if and only if $c_n \neq 0$. But $\det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \ldots + c_{n-1}\lambda + c_n$ or on setting $\lambda = 0$, we obtain $\det(-A) = c_n$ or $(-1)^n \det A = c_n$.

It follows from last equation that $\det A = 0$ if and only if $c_n = 0$, and this in turn implies that A is invertible if and only if $c_n \neq 0$.

Note : Using theorem (8), we see that the matrices in example (4.9) and (4.10) are **invertible**, since **none of these** matrices have eigenvalue $\lambda = 0$.

SOLVED EXAMPLES

Example 4.11 : Find det (A), given that A has $P(\lambda) = \lambda^3 - 2\lambda^2 + \lambda + 5$ as its characteristic polynomial.

Solution : Using theorem (8), we see that $\det(A) = (-1)^n c_n$, if A is an $n \times n$ matrix and c_n is the constant term in the characteristic polynomial of A.

In present case $c_3 = 5$, hence $\det(A) = (-1)^3 5 = -5$.

Example 4.12 : Find the eigen values and corresponding eigen vectors for each of the following matrices :

(i) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

(iii) $\begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$ (iv) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

Solution : (i) The characteristic equation is

$$0 = |A - \lambda I| = \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} = -\lambda^3 + 18\lambda^2 - 45\lambda$$

so that 0, 3, 15 are the three characteristic roots of the matrix.

If x, y, z be the components of a characteristic vector corresponding to the characteristic root, 0, we have

$$O = (A - 0I)\,X = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$\Rightarrow \quad 8x - 6y + 2z = 0,\ 6x + 7y - 4z = 0,\ 2x - 4y + 3z = 0$

These equations determine a single linearly independent solution which we may take as

$$[1 \quad 2 \quad 2]'$$

It may similar be shown by considering the equations

$$(A - 3I)\,X = 0, \quad (A - 15I)\,X = 0$$

that the characteristic vectors corresponding to the characteristic roots 3 and 15 are arbitrary non-zero multiples of the vectors

$$\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$$

The subspaces of V_3 spanned by these three vectors separately are the three characteristic spaces.

(ii) The characteristic equation is

$$0 = |A - \lambda I| = \begin{bmatrix} 6 - \lambda & 2 & 2 \\ -2 & 3 - \lambda & -1 \\ 2 & -1 & 3 - \lambda \end{bmatrix} = -\lambda^3 + 12\lambda^2 - 36\lambda + 32$$

so that 2, 2, 8 are the characteristic roots, only two roots are distinct.

Consider $(A - 8I)\,X = 0$, we may show that we obtain only one linearly independent solution

$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

so that every non-zero multiple of the same is a characteristic vector for the characteristic root 8.

For the characteristic root 2, we have

$$0 = (A - 2I)\,X = \begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$\Rightarrow \quad 4x - 2y + 2z = 0,\ -2x + y - z = 0,\ 2x - y + z = 0$

which are equivalent to a single equation.

Thus we obtain two linearly independent solutions which are

$$\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

The subspace of V_3 spanned by these two vectors is the characteristic space for the root 2.

(iii) The characteristic equation of the matrix is

$$-x^3 + 7x^2 - 16x + 12 = 0$$

so that the characteristic roots are 2, 2, 3.

corresponding to the characteristic root, 3, we find only one linearly independent characteristic vector which may be taken as

$$\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

For the repeated root, 2, we have

$$0 = (A - 2I)\, K$$

which gives

$$x + 10y + 5z = 0, \; -2x - 5z - 4z = 0, \; 3x + 5y + 5z = 0$$

These equations determine a single linearly independent solution.

This solution is $\begin{bmatrix} 5 \\ 2 \\ -5 \end{bmatrix}$

(iv) The characteristic equation is

$$(2 - x)^3 = 0$$

so that, 2, is the only distinct characteristic root.

It may be seen that $(A - 2I)\, X = 0$ determines only one linearly independent solution which is

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Example 4.13 : Show that a characteristic vector, X, corresponding to the characteristic root, λ, of a matrix A is also a characteristic vector of every matrix f(A); f(x) being any scalar polynomial, and the corresponding root for f(A) is $f(\lambda)$. In general, show that if

$$g(x) = f_1(x)/f_2(x); \; |f_2(A)| \neq 0$$

then $g(\lambda)$ is a characteristic root of

$$g(A) = f_1(A)\, \{f_2(A)\}^{-1}$$

Solution : If $AX = \lambda X$

then $A^2X = A(AX) = A(\lambda X) = \lambda(AX) = \lambda\lambda X = \lambda^2 X$

Repeating this process k times, we obtain

$$A^kX = \lambda^k X$$

$$\Rightarrow \quad f(A)\,X = (a_0I + a_1A + a_2A^2 + \ldots + a_mA^m)\,X$$

$$= a_0X + a_1\lambda X + \ldots + a_m\lambda^m X$$

$$= (a_0 + a_1\lambda + \ldots + a_m\lambda^m)\,X = f(\lambda)\,X$$

so that X is a characteristic vector of the matrix f(A) and $f(\lambda)$ is the corresponding characteristic root.

Since $|f_2(A)| \neq 0$, the matrix $f_2(A)$ is non-singular and any are characteristic root of $f_2(A)$ is not zero. In particular

$$f_2(\lambda) \neq 0$$

for $f_2(\lambda)$ is a characteristic root of $f_2(A)$. Now

$$f_1(A)\,X = f_1(\lambda)\,X \quad \ldots \text{(i)}$$

$$f_2(A)\,X = f_2(\lambda)\,X \quad \ldots \text{(ii)}$$

From (ii), $\{f_2(\lambda)\}^{-1}\,X = \{f(A)\}^{-1}\,X$

$$g(A)\,X = f_1(A)\,[\{f_2(A)\}^{-1}\,X] \quad \ldots \text{(iii)}$$

$$= f_1(A)\,\{f_2(\lambda)\}^{-1}\,X$$

$$= \{f_2(\lambda)\}^{-1}\,f_1(A)\,X$$

$$= \{f_2(\lambda)\}^{-1}\,f_1(\lambda)\,X = g(\lambda)\,X$$

Thus X is also a characteristic vector of g(A) with corresponding root (λ).

Example 4.14 : Show that the two matrices

$$A \cdot P^{-1}AP$$

the same characteristic roots.

Solution : Let

$$P^{-1}AP = B$$

$$\therefore \quad B - \lambda I = P^{-1}AP - \lambda I$$

$$= P^{-1}\,AP - P^{-1}\,\lambda IP = P^{-1}\,(A - \lambda I)\,P$$

$$\Rightarrow \quad |B - \lambda I| = |P^{-1}|\,|A - \lambda I|\,|P|$$

$$= |A - \lambda I|\,|P^{-1}|\,|P|$$

$$= |A - \lambda I|\,|P^{-1}P|$$

$$= |A - \lambda I|\,|I| = |A - \lambda I|$$

Thus the two matrices A and B have the same determinant hence the same characteristic equation and the same characteristic roots.

Another solution : $AX = \lambda X$

$\Rightarrow \quad P^{-1}AX = \lambda P^{-1}X$

$\Rightarrow \quad (P^{-1}AP)(P^{-1}X) = \lambda(P^{-1}X)$

that λ is also a characteristic root of $P^{-1}AP$ and $P^{-1}X$ is a corresponding characteristic vector.

Example 4.15 : If A is non-singular, prove that the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A.

Solution : Let λ be an eigen value of A and X be a corresponding eigen vector. Then

$$AX = \lambda X \Rightarrow X = A^{-1}(\lambda X) = \lambda(A^{-1}X)$$

$$\Rightarrow \frac{1}{\lambda}X = A^{-1}X \qquad [\because \text{ A is non-singular} \Rightarrow \lambda \neq 0]$$

$$\Rightarrow A^{-1}X = \frac{1}{\lambda}X$$

$\Rightarrow \frac{1}{\lambda}$ is an eigenvalue of A^{-1} and X is a corresponding eigenvector

Conversely, suppose that k is an eigenvalue of A^{-1}. Since A is non-singular $\Rightarrow A^{-1}$ is non-singular and $(A^{-1})^{-1} = A$, therefore it follows from the first part that $\frac{1}{k}$ is an eigenvalue of A. Thus each eigenvalue of A^{-1} is equal to the reciprocal of eigenvalue of A.

Example 4.16 : If α is a characteristic root of a non-signular matrix A, then prove that $\frac{|A|}{\alpha}$ is a characteristic root of Adj. A.

Solution : Since α is a characteristic root of a non-singular matrix, therefore $\alpha \neq 0$. Also α is a characteristic root of A implies that there exists a non-zero vector X such that

$$AX = \alpha X$$

$\Rightarrow$ (Adj. A) (AX) = (Adj. A) (αX)

$\Rightarrow$ [(Adj. A) A] X = α (Adj. A) X

$\Rightarrow$ |A| IX = α (Adj. A) X $\qquad [\because$ (Adj. A) A = |A| I]

$\Rightarrow$ |A| X = α (Adj. A) X

$\Rightarrow \frac{|A|}{\alpha}X = (\text{Adj. A})X$

$\Rightarrow (\text{Adj. A})X = \frac{|A|}{\alpha}X$

Since X is a non-zero vector, therefore $\frac{|A|}{\alpha}$ is a characteristic root of the matrix Adj. A.

Exercise 4.1

1. Find out the eigenvalues and the corresponding eigenspaces of the following matrices :

(i) $\begin{bmatrix} 3 & -5 & -4 \\ -5 & -6 & -5 \\ -4 & -5 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 5 & 4 & -5 \\ 4 & 5 & -4 \\ -1 & -1 & 2 \end{bmatrix}$

(iii) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (iv) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

(v) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \end{bmatrix}$ (iv) $\begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

(vii) $\begin{bmatrix} 8 & -3 & -6 & -6 \\ 4 & -1 & -4 & -4 \\ 3 & -5 & 1 & 1 \\ 2 & 2 & -4 & -4 \end{bmatrix}$

2. Suppose $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$ find the eigenvalues of A and the corresponding, eigen vectors : Verify that these eigen vectors are pairwise orthogonal.

3. Find the characteristic values and the characteristic vectors of the matrix :

$$\begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

4. Check whether matrix $A = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ is orthogonal matrix.

5. Find the eigen values and the corresponding eigen vectors of the matrix B

$$B = \begin{bmatrix} 5 & 3 & 3 \\ 3 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix}$$

Construct a matrix C such that $C^{-1}BC = D$ where D is a diagonal matrix.

Write down the matrices C^{-1} and D.

6. Find the characteristic roots and the characteristic vectors of the matrix :

$$\begin{bmatrix} 2 & -3 & 1 \\ -3 & 1 & 3 \\ -5 & 2 & -4 \end{bmatrix}$$

7. Show that the matrices A and A' have the same eigenvalues.
8. Show that 0 is a characteristic root of a matrix if, and only if the matrix is singular.
9. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A, then show that $k\lambda_1, \ldots, k\lambda_n$ are the eigenvalues of kA.
10. Show that if λ is a characteristic root of the matrix A then $k + \lambda$ is a characteristic root of the matrix $A + kI$.
11. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the characteristic roots of the n-square matrix A and k is a scalar, prove that characteristic roots of $A - kI$ are $\alpha_1 - k, \alpha_2 - k, \ldots, \alpha_n - k$.

Theorem 10 : Any two characteristic vectors corresponding to two distinct characteristic roots of a Hermitian matrix are orthogonal.

Let X_1, X_2 be two characteristic vectors corresponding to two distinct characteristic roots λ_1, λ_2 of a Hermitian matrix A. We have

$$AX_1 = \lambda_1 X_1 \quad \ldots (1)$$

$$AX_2 = \lambda_2 X_2 \quad \ldots (2)$$

where λ_1, λ_2 are real numbers.

Pre-multiplying with X_2^θ, X_1^θ respectively, we obtain

$$X_2^\theta AX_1 = \lambda_1 X_2^\theta X_1$$

$$X_1^\theta AX_2 = \lambda_2 X_1^\theta X_2$$

But $(X_2^\theta AX_1)^\theta = X_1^\theta AX_2$, for $A^\theta = A$,

$\therefore$ $(\lambda_1 X_2^\theta X_1)^\theta = \lambda_2 X_1^\theta X_2$

$\Rightarrow$ $\lambda_1 X_1^\theta X_2 = \lambda_2 X_2^\theta X_2$

$\Rightarrow$ $(\lambda_1 - \lambda_2) X_1^\theta X_2 = 0$

$\therefore$ $X_1^\theta X_2 = 0$, for $\lambda_1 - \lambda_2 \neq 0$

Thus the vectors X_1, X_2 are orthogonal.

Corollary : Any two characteristic vectors corresponding to two distinct characteristic rots of a real symmetric matrix are orthogonal.

Theorem 11 : Any two characteristic vectors corresponding to two distinct characteristic roots of a unitary matrix are orthogonal.

Let $AX_1 = \lambda_1 X_1$ … (1)

$AX_2 = \lambda_2 X_2$ … (2)

where $\lambda_1 \neq \lambda_2$.

Taking conjugate transpose of (2).

$X_1^\theta A^\theta = \lambda_1 X_2^\theta$ … (3)

From equation (1)and (3),

$X_2^\theta A^\theta A X_1 = \lambda_2 \lambda_1 X_2^\theta X_1$

$\Rightarrow \quad (1 - \lambda_2\lambda_1) X_2^\theta X_2 = 0$, for $A^\theta A = I$ … (4)

As, A, is a unitary matrix, the modulus of each of its characteristics roots is unity so that

$\lambda_2\lambda_2 = 1$

Thus, $(1 - \lambda_2\lambda_1) = \lambda_2\bar{\lambda}_1 - \bar{\mu}_1\lambda_2 = \bar{\lambda}_2 (\bar{\lambda}_2 - \lambda_1) \neq 0$ … (5)

From (4) and (5), we deduce that

$X_2^\theta X_1 = 0$

i.e., X_1 and X_2 are orthogonal.

SOLVED EXAMPLES

Example 4.17 : If A is both real symmetric and orthogonal, prove that all its eigenvalues are + 1 or – 1.

Solution : If A is a real symmetric matrix, than all its eigenvalues are real. Further, if A is orthogonal, then all its eigenvalues must be of unit modulus. Now ± 1 are the only real numbers of unit modulus. Hence, if A is both real symmetric and orthogonal, then all its eigenvalues are + 1 or – 1.

Example 4.18 : Find the characteristic roots of the 2 × 2 orthogonal matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Solution : We have,

$$|A - \lambda I| = \begin{bmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{bmatrix} = (\cos\theta - \lambda)^2 + \sin^2\theta$$

$\therefore$ Characteristic equation of A is

$(\cos\theta - \lambda)^2 + \sin^2\theta = 0$

$\Rightarrow \quad \cos\theta - \lambda = \pm i \sin\theta$

$\Rightarrow \quad \lambda = \cos\theta \pm i \sin\theta$

which are of unit modulus.

Exercises 4.2

1. A real skew-symmetric matrix S satisfies the relation $A^2 + I = 0$ where I is the identity matrix. Show that A is orthogonal and even order matrix.
2. A matrix A satisfies the relation $A^2 = A$ and is neither identity nor the null matrix. Show that the characteristic equation of A is of the form $\lambda^p(1-\lambda)^q = 0$ and that A is neither orthogonal nor skew-symmetric.
3. Prove that ± 1 can be the only real characteristic roots of an orthogonal matrix.
4. Show that the roots of the equation $\begin{vmatrix} a+x & h & g \\ h & b+x & f \\ g & f & c+x \end{vmatrix} = 0$ are real; a, b, c, f, g, h being real numbers.
5. Find the characteristic roots and characteristic vectors of the matrix

$$A = \begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$$

Also verify that the geometric multiplicity of a characteristic root cannot exceed its algebraic multiplicity.

4.4 The Construction of Orthogonal Matrices

The following theorem gives a very simple method for constructing orthogonal matrices and was discovered by Cayley in 1846.

Theorem 12 : If S, is a real skew-symmetric matrix, then I – S is non-singular and the matrix.

$$A = (I + S)(I - S)^{-1}$$

is orthogonal.

Firstly, we show that I – S is non-singular.

The equality $|I - S| = 0$

implies that I is a characteristic root of the matrix S, but this is not possible, for a real skew-symmetric matrix can have only zero or pure imaginary numbers as its characteristic roots. Thus

$$|I - S| \neq 0$$

i.e., I – S is non-singular.

We have $A' = [(I - S)^{-1}]'(I + S)' = [(I - S)']^{-1}(I + S)'$

But $(I - S)' = I' - S' = I + S'$

$(I + S)' = I' + S' = I - S$

$\therefore \quad A' = (I + S)^{-1}(I - S)$

$\therefore \quad A'A = (I + S)^{-1}(I - S)(I + S)(I - S)^{-1}$

$= (I + S)^{-1}(I + S)(I - S)(I - S)^{-1} = 1$

Example 4.19 : If A be an orthogonal matrix with the property that − I is not a characteristic root, then A is expressible as

$$(I + S)\,(I - S)^{-1}$$

for some suitable real skew symmetric matrix S.

Solution : The problem will be solved if we show that corresponding to a given orthogonal matrix A, such that − 1 is not a characteristic root of A, the equation

$$A = (I + S)\,(I - S)^{-1} \quad \ldots (1)$$

is solvable for S and the solution is a skew-symmetric matrix.

Equation (i), $\Rightarrow A\,(I - S) = (I + S)$

$\Rightarrow A - AS = I + S$

$\Rightarrow A - I = AS + S$

$\Rightarrow (A - I) = (A + I)\,S \quad \ldots (2)$

Since − 1 is not a characteristic root of A, therefore $|A + I| \neq 0$, i.e. (A + I) is non-singular. Therefore, pre-multiplying both sides of (1) by $(A + I)^{-1}$, we get

$$(A + I)^{-1}\,(A - I) = S$$

Thus (1) is solvable for S. Since A is a real matrix, therefore S is also a real matrix. Now it remains to show that S is a skew-symmetric matrix. We have

$$\begin{aligned} S' &= [(A + I)^{-1}\,(A - I)]' \\ &= (A - I)'\,[(A + I)^{-1}]' \\ &= (A - I)'\,[(A + I)']^{-1} = (A' - I')\,[(A' + I')]^{-1} \\ &= (A' - I')\,(A' + I)^{-1} \quad \ldots (3) \end{aligned}$$

It can be easily seen that (A' − I) and (A' + I) commute. Therefore

$(A' + I)\,(A' - I) = (A' - I)\,(A' + I)$

$\Rightarrow (A' + I)^{-1}\,(A' + I)\,(A' - I)\,(A' + I)^{-1} = (A' + I)^{-1}\,(A' - I)\,(A' + I)\,(A' + I)^{-1}$

$\Rightarrow (A' - I)\,(A' + I)^{-1} = (A' + I)^{-1}\,(A' - I)$

From (3), we get

$$\begin{aligned} S' &= (A' + I)^{-1}\,(A' - I) = (A' + A'A)^{-1}\,(A - A'A) \\ &\qquad [\because A \text{ is orthogonal} \Rightarrow A'A + I] \\ &= [A'\,(I + A)]^{-1}\,[A'\,(I - A)] \\ &= (I + A)^{-1}\,(A')\,A'\,(I - A) \\ &= (I + A)^{-1}\,(I - A) = (A + I)^{-1}\,(I - A) = -S \end{aligned}$$

$\therefore$ S is a skew-symmetric matrix.

4.5 Diagonalization of Matrices

In this section, we shall be concerned with the problem of finding a basis of $\mathbb{R}^n$ that consists of eigenvectors of a given n × n matrix A. In this connection, we have the following two problems, which are actually equivalent.

The Eigenvector Problem : Given an n × n matrix A, does there exist a basis for $\mathbb{R}^n$ consisting of eigenvectors of A ?

The Diagonalization Problem (Matrix form) : Given an $n \times n$ matrix A, does there exist an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix ?

The following theorem proves the equivalence of the above two problems, before that we define diagonalizable matrix.

Definition : A square matrix A is called *diagonalizable* if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix, the matrix P is said to diagonalize A.

Theorem 13 : If A is an $n \times n$ matrix, then the following are equivalent :

(a) A is diagonalizable.

(b) A has n linearly independent eigenvectors.

Proof : We prove the theorem by establishing (a) $\Rightarrow$ (b) and (b) $\Rightarrow$ (a).

To prove (a) $\Rightarrow$ (b).

Since A is assumed diagonalizable, there is an invertible matrix

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdots\cdots & P_{1n} \\ P_{21} & P_{22} & \cdots\cdots & P_{2n} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & \cdots\cdots & P_{nn} \end{bmatrix}$$

such that, $P^{-1}AP$ is diagonal, say $P^{-1}AP = D$, where

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots\cdots & 0 \\ 0 & \lambda_2 & \cdots\cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots\cdots & \lambda_n \end{bmatrix}$$

It follows from the hypothesis $P^{-1}AP = D$ that $AP = PD$; that is,

$$AP = \begin{bmatrix} P_{11} & P_{12} & \cdots\cdots & P_{1n} \\ P_{21} & P_{22} & \cdots\cdots & P_{2n} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & \cdots\cdots & P_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots\cdots & 0 \\ 0 & \lambda_2 & \cdots\cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots\cdots & \lambda_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 P_{11} & \lambda_2 P_{12} & \cdots\cdots & \lambda_n P_{1n} \\ \lambda_1 P_{21} & \lambda_2 P_{22} & \cdots\cdots & \lambda_n P_{2n} \\ \vdots & & & \vdots \\ \lambda_1 P_{n1} & \lambda_2 P_{n2} & \cdots\cdots & \lambda_n P_{nn} \end{bmatrix} \quad \ldots(1)$$

If $P_1, P_2, \ldots, P_n$ denote the column vectors of P, then from (1), the successive columns on R.H.S. of A_1 are $\lambda_1 P_1, \lambda_2 P_2, \ldots, \lambda_n P_n$. However, the successive columns of AP are $AP_1, AP_2, \ldots, AP_n$. Thus, we must have,

$$AP_1 = \lambda_1 P_1, AP_2 = \lambda_2 P_1, \ldots, AP_n = \lambda_n P_n \quad \ldots(2)$$

Since P is invertible, its column vectors are all non-zero ; thus, it follows from (2) that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A, and $P_1, P_2, \ldots, P_n$ are corresponding eigenvectors. Since P is invertible, it follows that $P_1, P_2, \ldots, P_n$ are linearly independent. Thus, A has n linearly independent eigenvectors.

To prove (b) $\Rightarrow$ (a)

Assume that A has n linearly independent eigenvectors $P_1, P_2, \ldots, P_n$, with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and let

$$P = \begin{bmatrix} P_{11} & P_{12} & \ldots\ldots\ldots & P_{1n} \\ P_{21} & P_{22} & \ldots\ldots\ldots & P_{2n} \\ : & : & & : \\ P_{n1} & P_{n2} & \ldots\ldots\ldots & P_{nn} \end{bmatrix}$$

be the matrix whose column vectors are $P_1, P_2, \ldots\ldots, P_n$. Again we know that the columns of the product AP are,

$$AP_1, AP_2, \ldots\ldots\ldots, AP_n$$

But $AP_1 = \lambda_1 P_1, AP_2 = \lambda_2 P_2, \ldots\ldots\ldots, AP_n = \lambda_n P_n$ so that,

$$AP = \begin{bmatrix} \lambda_1 P_{11} & \lambda_2 P_{12} & \ldots\ldots\ldots & \lambda_n P_{1n} \\ \lambda_1 P_{21} & \lambda_2 P_{22} & \ldots\ldots\ldots & \lambda_n P_{2n} \\ : & : & & : \\ \lambda_1 P_{n1} & \lambda_2 P_{n2} & \ldots\ldots\ldots & \lambda_n P_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} P_{11} & P_{12} & \ldots\ldots\ldots & P_{1n} \\ P_{21} & P_{22} & \ldots\ldots\ldots & P_{2n} \\ : & : & & : \\ P_{n1} & P_{n2} & \ldots\ldots\ldots & P_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots\ldots\ldots & 0 \\ 0 & \lambda_2 & \ldots\ldots\ldots & 0 \\ : & : & & : \\ 0 & 0 & \ldots\ldots\ldots & \lambda_n \end{bmatrix}$$

$$= PD \quad \ldots(3)$$

where, D is the diagonal matrix having the eigenvalues $\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_n$ on the main diagonal. Since, the column vectors of P are linearly independent, P is invertible; thus, equation (3) can be rewritten as $P^{-1}AP = D$; that is, A is diagonalizable. This completes the proof of the theorem.

The step-by-step procedure to check whether given matrix is diagonalizable :

Let A be an $n \times n$ matrix, then the following are the steps to check whether A is diagonalizable.

Step 1 : Form the matrix $\lambda I - A$.

Step 2 : Find the characteristics polynomial $\Delta(\lambda) = \det(\lambda I - A)$ of A.

Step 3 : Solve the equation $\Delta(\lambda) = \det(\lambda I - A) = 0$ for λ, to determine eigenvalues of A. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A.

Step 4 : For each eigenvalue λ_i, i = 1, 2, …, n find eigenspace $E(\lambda_i)$ and its basis.

Step 5 : The Totality of the basis vectors of each eigenspace $E(\lambda_i)$ obtained in step 4 are the eigenvectors of A. The total number of basis vectors so obtained will be $\leq n$.

Step 6 : If there are n linearly independent eigenvectors of A obtained in step 5, then A is diagonalizable and the matrix P that diagonalizes A is the matrix whose columns are basis vectors found in step 5.

Step 7 : if there are less than n linearly independent eigenvectors found in step 5, then A is not diagonalizable.

FLOW CHART

To determine given n × n matrix A is diagonalizable.

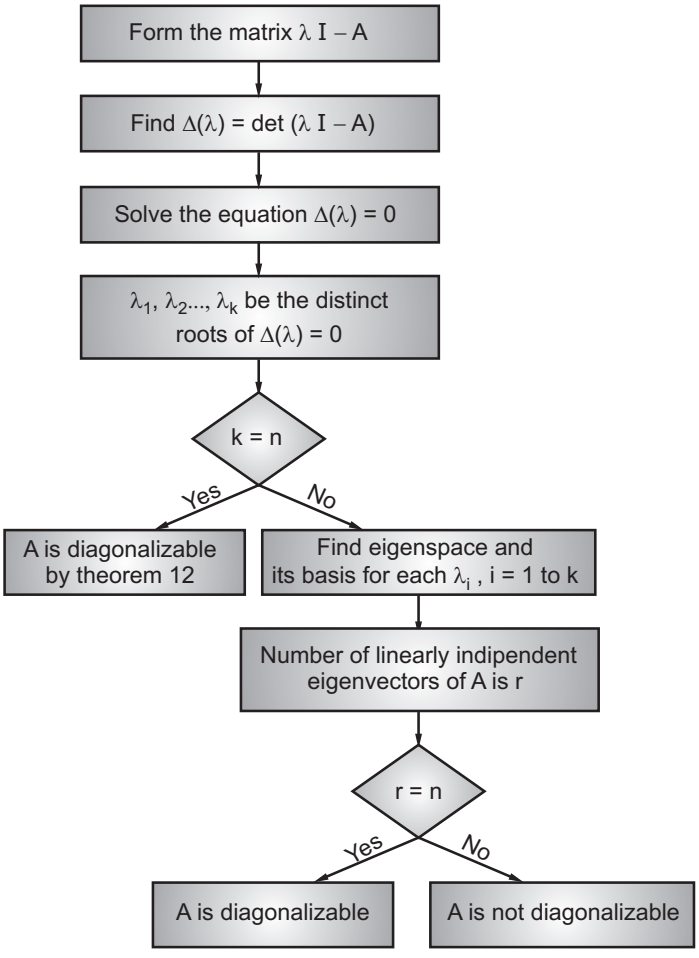

Form the matrix λ I – A
Find Δ(λ) = det (λ I – A)
Solve the equation Δ(λ) = 0
λ1, λ2..., λk be the distinct roots of Δ(λ) = 0
k = n
Yes
No
A is diagonalizable by theorem 12
Find eigenspace and its basis for each λi , i = 1 to k
Number of linearly indipendent eigenvectors of A is r
r = n
Yes
No
A is diagonalizable
A is not diagonalizable

Fig. 4.1

SOLVED EXAMPLES

Example 4.20 : Find a matrix P that diagonalizes

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

Solution : From example 4.5 of the previous section, the eigenvalues of A are $\lambda = 1, -1, 3$. Also from that example, the vectors

$$P_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; P_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \text{ and } P_3 = \begin{bmatrix} -2 \\ -3 \\ 1 \end{bmatrix}$$

are the eigenvectors corresponding to the eigenvalues 1, –1, 3 respectively.

Also P_1 forms a basis for E ($\lambda = 1$),

P_2 forms a basis for E ($\lambda = -1$) and

P_3 forms a basis for E ($\lambda = 3$).

It is easy to verify that $\{P_1, P_2, P_3\}$ is linearly independent, so that,

$$P = \begin{matrix} P_1 & P_2 & P_3 \\ \downarrow & \downarrow & \downarrow \end{matrix} \begin{bmatrix} -1 & 0 & -2 \\ 1 & -1 & -3 \\ 0 & 1 & 1 \end{bmatrix}$$

diagonalizes A.

Problem : The students are requested to verify that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Example 4.21 : Show that the matrix,

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$$

is not diagonalizable.

Solution : The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & 0 \\ -1 & \lambda - 2 \end{bmatrix} = \lambda^2 - 2\lambda + 4 = (\lambda - 2)^2$$

Therefore, the eigenvalues of A are given by,

$$\det(\lambda I - A) = 0$$

or

$$(\lambda - 2)^2 = 0$$

Thus, $\lambda = 2$, is the only eigenvalue of A. Let us find eigenvector of A corresponding to $\lambda = 2$. For this, we have to solve the equation

$$\begin{bmatrix} \lambda - 2 & 3 \\ -1 & \lambda - 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \ldots(*)$$

For $\lambda = 2$, $(*)$ becomes,

$$\begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In this system x_2 is free variable, the solution is given by $x_1 = 0$, $x_2 = t$, $t \in \mathbb{R}$.

$$\therefore \quad E(\lambda = 2) = \left\{ (0, t) \,\middle|\, t \in \mathbb{R} \right\} = \left\{ t(0, 1) \,\middle|\, t \in \mathbb{R} \right\}$$

Since, this space is one-dimensional. A does not have two linearly independent eigenvectors and is therefore, **not diagonalizable**.

Example 4.22 : Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$$

In example (4.6) of section 4.2 we have seen that $\lambda = 2, 2, 4$ are eigenvalues of A. Further, we have obtained : The eigenspace for $\lambda = 2$ to be $E(\lambda = 2) = \{t(-1, 1, 1) | t \in \mathbb{R}\}$ and the eigenspace for $\lambda = 4$ is

$$E(\lambda = 4) = \{t(1, -1, 1) \mid t \in \mathbb{R}).$$

Thus we see that there are only two linearly independent eigenvectors, namely $(-1, 1, 1)$ and $(1, -1, 1)$ for A, hence A is *not diagonalizable*.

Definition : Let A and B be square matrices such that $B = P^{-1}AP$ for an invertible matrix P, then *B is said to be similar to A.*

Theorem 14 : Similar matrices have the same characteristic polynomial.

Proof : Let A and B be two similar matrices. Then there is an invertible matrix P such that $B = P^{-1}AP$. Using $\lambda I = P^{-1}(\lambda I) P$, we have,

$$\begin{aligned} \det(\lambda I - B) &= \det(\lambda I - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I) P - P^{-1}AP) \\ &= \det[P^{-1}(\lambda I - A) P] \\ &= \det P^{-1} \cdot \det(\lambda I - A) \det P \\ &= \det(\lambda I - A). \text{ Since } \det P^{-1} \cdot \det P = 1. \end{aligned}$$

Thus, the characteristic polynomials of A and B are same.

Corollary : Similar matrices have the same eigenvalues.

Theorem 15 : If $\vec{v}_1, \vec{v}_2 \ldots\ldots \vec{v}_r$ are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_r$ then $\{\vec{v}_1, \vec{v}_2, \ldots\ldots, \vec{v}_r\}$ is linearly independent set.

Proof : Let us prove this theorem by mathematical induction on r.

For r = 1, $\vec{v_1}$ is eigenvector corresponding to the eigenvalue λ_1 of A, since eigenvector is non-zero, by definition $\{\vec{v_1}\}$ is linearly independent, hence, the result is true for r = 1. Now, let us assume that the result is true for r – 1 eigenvectors ; that is, any r – 1, eigenvectors of A corresponding to distinct r – 1 eigenvalues are linearly independent. Now, suppose $\vec{v_1}$, $\vec{v_2}$,……, $\vec{v_r}$ be r eigenvectors of A corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ of A.

Now, suppose we have,

$$\alpha_1\vec{v_1} + \alpha_2\vec{v_2} + \ldots\ldots + \alpha_r\vec{v_r} = 0 \qquad \ldots(1)$$

Multiplying both sides of (1) by A and using $A\vec{v_1} = \lambda_1\vec{v_1}$, $A\vec{v_2} = \lambda_2\vec{v_2}$, … $A\vec{v_r} = \lambda_r\vec{v_r}$, we obtain

$$\alpha_1\lambda_1\vec{v_1} + \alpha_2\lambda_2\vec{v_2} + \ldots\ldots + \alpha_r\lambda_r\vec{v_r} = \vec{0} \qquad \ldots(2)$$

Multiplying both sides of (1) by λ_r and subtracting the resulting equation from (2) yields

$$\alpha_1(\lambda_1 - \lambda_r)\vec{v_1} + \alpha_2(\lambda_2 - \lambda_r)\vec{v_2} + \ldots\ldots + \alpha_{r-1}(\lambda_{r-1} - \lambda_r)\vec{v_{r-1}} = \vec{0}$$

Since $\{\vec{v_1}, \vec{v_2}, \ldots\ldots, \vec{v_{r-1}}\}$ is linearly independent set, by induction hypothesis, this equation implies that,

$\alpha_1(\lambda_1 - \lambda_r) = \alpha_2(\lambda_2 - \lambda_r) = \ldots\ldots = \alpha_{r-1}(\lambda_{r-1} - \lambda_r) = 0$ and since $\lambda_1, \lambda_2, \ldots\ldots, \lambda_r$ are distinct, $\lambda_i - \lambda_r \neq 0$, it follows that,

$$\alpha_1 = \alpha_2 = \ldots\ldots\ldots\ldots = \alpha_{r-1} = 0.$$

Substituting these values in (1) yields

$$\alpha_r\vec{v_r} = 0$$

Since the eigenvector $\vec{v_r}$ is non-zero, it follows that $\alpha_r = 0$.

Thus, the equation (1) implies that

$$\alpha_1 = \alpha_2 = \ldots\ldots\ldots = \alpha_r = 0.$$

Showing that $\{\vec{v_1}, \vec{v_2}, \ldots\ldots \vec{v_r}\}$ is linearly independent.

This completes the proof.

The consequence of this theorem is the following significant result.

Theorem 16 : If A is n × n matrix and A has n distinct eigenvalues, then A is diagonalizable.

Proof : If $\vec{v_1}$, $\vec{v_2}$, ……, $\vec{v_n}$ are eigenvectors corresponding to the eigenvalues λ_1, λ_2, ……, λ_n, then by theorem 11, $\vec{v_1}$, $\vec{v_2}$, …… $\vec{v_n}$ are linearly independent. Thus, A has n linearly independent eigenvectors, hence A is diagonalizable by theorem 9.

SOLVED EXAMPLES

Example 4.23 : We saw in example (4.3) in the previous section that,

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

has three distinct eigenvalues $\lambda = 1, -1, 3$. Therefore, A is diagonalizable as a consequence of theorem 12. Further,

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

for some invertible matrix P.

Example 4.24 : The converse of the theorem 12 is not true; that is, an n × n matrix A may be diagonalizable even if it does not have n distinct eigenvalues. For example, (a) the only eigenvalue of the diagonal matrix

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

is $\lambda = 2$, yet A is obviously diagonalizable by taking P = I, since,

$$P^{-1}AP = I^{-1}AI = A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

(b) In example 4.7 of the previous section, we have seen that,

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

has eigenvalues $\lambda = 1$ and $\lambda = 7$, not distinct. However, the eigenvectors $P_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $P_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ of A corresponding to $\lambda = 1$, are linearly independent. Further $P_3 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ 0 \end{bmatrix}$ is the eigenvector of A corresponding to $\lambda = 7$. It can be seen that $\{P_1, P_2, P_3\}$ is linearly independent (students should prove this as an exercise). Thus, A has three linearly independent eigenvectors, hence, by theorem 12, A is diagonalizable.

Thus, A is diagonalizable without having distinct eigenvalues.

Example 4.25 : Determine a, b, c, d, e and f, given that $\vec{v_1} = (1, 1, 1)$, $\vec{v_2} = (1, 0, -1)$ and $\vec{v_3} = (1, -1, 0)$ are eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}$$

Solution : Suppose λ_1, λ_2 and λ_3 be the eigenvalues of A corresponding to the eigenvectors v_1, v_2 and v_3 respectively. Then, by definition,

$$A\vec{v_1} = \lambda_1\vec{v_1}, A\vec{v_2} = \lambda_2\vec{v_2}, \text{ and } A\vec{v_3} = \lambda_3\vec{v_3}.$$

or $$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \lambda_1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \lambda_2\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

and $$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \lambda_3\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Simplifying and equating corresponding entries of the above three equations, we obtain nine equations :

$$\begin{aligned} 3 &= \lambda_1 \\ a + b + c &= \lambda_1 \\ d + e + f &= \lambda_1 \\ 0 &= \lambda_2 \\ a + 0 - c &= \lambda_2 \cdot 0 = 0 \\ d - f &= (-1)\,\lambda_2 \\ 0 &= \lambda_3 \\ a - b &= \lambda_3\,(-1) \\ d - e &= \lambda_3\,(0) = 0 \end{aligned}$$

Solving these equations, we obtain,

$\lambda_1 = 3, \lambda_2 = 0, \lambda_3 = 0, a = b = c = d = e = f = 1.$

Geometric and Algebraic Multiplicity :

Definition : Let λ_0 be an eigenvalue of an $n \times n$ matrix A, then the dimension of the eigenspace corresponding to λ_0 is called the **geometric multiplicity of λ_0** and the number of times that $(\lambda - \lambda_0)$ appears as a factor in the characteristics polynomial of A is called the **algebraic multiplicity** of A.

SOLVED EXAMPLES

Example 4.26 : Find geometric and algebraic multiplicity of each of the eigenvalues of

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

Solution : We have seen in example (4.7) in the previous section that the characteristic polynomial of A is

$$\begin{aligned} \det(\lambda I - A) &= \lambda^3 - 9\lambda^2 + 15\lambda - 7 \\ &= (\lambda - 1)^2\,(\lambda - 7) \end{aligned}$$

Thus, $\lambda = 1$ and $\lambda = 7$ are eigenvalues of A, from the characteristic polynomial we see that **algebraic multiplicity of $\lambda = 1$ is 2** and the **algebraic multiplicity of $\lambda = 7$ is 1**.

We calculated the eigenspace $E(\lambda = 1)$ in example 4.7 to be

$$\begin{aligned} E(\lambda = 1) &= \{(-r - s, r, s) \mid r, s \in \mathbb{R}\} \\ &= \{r(-1, 1, 0) + s(-1, 0, 1) \mid r, s \in \mathbb{R}\} \end{aligned}$$

Thus, $\dim (E(\lambda = 1)) = 2$

Therefore, the **geometric** multiplicity of $\lambda = 1$ is 2.

Also, it is calculated in example (4.7) that the eigenspace corresponding $\lambda = 7$ is given by

$$E(\lambda = 7) = \left\{\left(\frac{1}{3}t, -\frac{2}{3}t, t\right) \mid t \in \mathbb{R}\right\}$$

$$= \left\{t\left(\frac{1}{3}, -\frac{2}{3}, 1\right) \mid t \in \mathbb{R}\right\}$$

So that $\dim E(\lambda = 7) = 1$, hence the **geometric multiplicity of $\lambda = 7$ is 1**.

We state below the theorem without proof, which signifies the importance of geometric and algebraic multiplicity of an eigenvalues.

Theorem 17 : If A is a square matrix, then

(a) For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity.

(b) A is diagonalizable if and only if, for every eigenvalue, the geometric multiplicity is equal to algebraic multiplicity.

Remark : Students are requested to verify the theorem for the matrices in examples (4.4), (4.5), (4.6) and (4.10).

Computing power of a matrix :

In applied mathematics, one requires to compute higher powers of a square matrix. The next theorem shows how diagonalization can be used to simplify such computations for diagonalizable matrices.

Theorem 18 : Let A be a diagonalizable matrix and let P be the invertible matrix that diagonalizes A, then

$$A^k = P D^k P^{-1}$$

where, $P^{-1}AP = D$, a diagonal matrix.

Proof : Since A is diagonalized by P, we have

$$P^{-1}AP = D, \text{ a diagonal matrix.}$$

We first note that $(P^{-1}AP)^k = P^{-1}A^kP$, for any positive integer k. This can be proved by mathematical induction.

Therefore, $(P^{-1}AP)^k = D^k$ for any positive integer k.

That is $P^{-1}A^kP = D^k$

Solving this for A^k, we get

$$A^k = PD^kP^{-1}$$

Note : The above theorem expresses k-th power of A in terms of the k-th power of the diagonal matrix D. Obviously D^k is easy to compute. For if

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ : & : & & : \\ 0 & 0 & & d_n \end{bmatrix}$$

then $$D^k = \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ : & : & & : \\ 0 & 0 & \dots & d_n^k \end{bmatrix}$$

Example 4.27 : Compute A^{17}, where $A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$.

Solution : From example (4.12), we know that A is diagonalizable and the matrix P that diagonalises A is given by

$$P = \begin{bmatrix} -1 & 0 & -2 \\ 1 & -1 & -3 \\ 0 & 1 & 1 \end{bmatrix}$$

Now we have, $$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = D$$

$\therefore$ $$A^{17} = PD^{17}P^{-1} \qquad \dots (*)$$

We have, $$D^{17} = \begin{bmatrix} 1^{17} & 0 & 0 \\ 0 & (-1)^{17} & 0 \\ 0 & 0 & 3^{17} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3^{17} \end{bmatrix}$$

$$P^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{5}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix}$$

Now,
$$PD^{17}P^{-1} = \begin{bmatrix} -1 & 0 & -2 \\ 1 & -1 & -3 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3^{17} \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{5}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}(1+3^{17}) & -\frac{1}{2}(1-3^{17}) & -\frac{1}{2}(1-3^{17}) \\ -\frac{1}{4}(1-3^{18}) & \frac{1}{4}(3+3^{18}) & \frac{1}{4}(7+3^{18}) \\ -\frac{1}{4}(1+3^{17}) & -\frac{1}{4}(1+3^{17}) & -\frac{1}{4}(5+3^{17}) \end{bmatrix}$$

4.6 Minimal Polynomial and Minimal Equation of a Matrix

Let f(x) is a polynomial in the form of x and A is a square matrix of order n. If f(A) = 0, then we say that the polynomial f(x) is satisfied by the matrix A. Every matrix satisfies its characteristic equation and the characteristic polynomial of a matrix A is a non-zero polynomial, i.e,, a polynomial in which the coefficients of various terms are not all zero. Therefore, at least the characteristic polynomial of A is a non-zero polynomial that satisfied by matrix A. Thus the set of those non-zero polynomials which are satisfied by matrix A empty is not satisfied by matrix.

Monic Polynomial : A polynomial of x in which the coefficient of the highest power is unity is called a monic polynomial, e.g., $x^3 + 4x^2 - 3x + 5$ is a monic polynomial of degree 3 over the field of real numbers.

Among those non-zero polynomials which are satisfied by matrix A, the polynomial which is monic and which is of the lowest degree is of special interest. It is called the *minimal polynomial* of the matrix A.

Minimal Polynomial of a Matrix : The monic polynomial of lowest degree that is satisfied by a matrix A is called the minimal polynomial of A and if f(x) is the minimal polynomial of A, the equation f(x) = 0 is called the minimal equation of the matrix A.

If A is of order n, then its characteristic polynomial is of degree n. Since the characteristic polynomial of A annihilates A, therefore the minimal polynomial of A cannot be of degree greater than n. Its degree must be less than or equal to n.

Theorem 20 : The minimal polynomial of a matrix is unique.

Let the minimal polynomial of a matrix A is of degree r. Then no non-zero polynomial of degree less than r can annihilate A. Let

$$f(x) = x^r + a_1x^{r-1} + a_2x^{r-2} + \ldots + a_{r-1}x + a_r$$

and
$$g(x) = x^r + b_1x^{r-1} + b_2x^{r-2} + \ldots + b_{r-1}x + b_r$$

be two minimal polynomials of A. Then both f(x) and g(x) annihilate A.

Therefore we have, $f(A) = 0$ and $g(A) = 0$. These give

$$A^r + a_1A^{r-1} + \ldots + a_{r-1}A + a_rI = 0 \quad \ldots(1)$$

and $$A^r + b_1A^{r-1} + \ldots + b_{r-1}A + b_1I = 0 \quad \ldots (2)$$

Subtracting equation (1) from (2), we get

$$(b_1 - a_1)\, A^{r-1} + \ldots + (b_r - a_r)\, I = 0 \quad \ldots (3)$$

From (3), we see that the polynomial on L.H.S. also annihilates this matrix.

Let $m(x)$ be the minimal polynomial of a matrix A. Let $h(x)$ be any polynomial that annihilates A. Since $m(x)$ and $h(x)$ are two polynomials, hence by the **division algorithm**

$$h(x) = m(x)\, q(x) + r(x) \quad \ldots (1)$$

where either $r(x)$ is a zero polynomial or its degree is less than the degree of $m(x)$. Putting $x = A$ on both sides of (1), we get

$$h(A) = m(A)\, q(A) + r(A)$$

$$\Rightarrow \quad 0 = 0 \cdot q(A) + r(A) \qquad [\because \text{ both } m(x) \text{ and } h(x) \text{ annihilate A}]$$

$$\Rightarrow \quad r(A) = 0$$

Thus $r(x)$ is polynomial which also annihilates A. If $r(x) \neq 0$, then it is a non-zero polynomial of degree smaller than the degree of minimal polynomial $m(x)$ and thus we arrive at a contradiction that $m(x)$ is minimal polynomial of A. Therefore $r(x)$ must be a zero polynomial. Then (1) gives

$$h(x) = m(x)\, q(x) \Rightarrow m(x) \text{ is a divisor of } h(x)$$

Cor. 1 : The minimal polynomial of a matrix is a divisor of the characteristic polynomial of that matrix.

Let $f(x)$ be the characteristic polynomial of a matrix A. Then $f(A) = 0$.

Cayley-Hamilton theorem. Thus $f(x)$ annihilates A. If $m(x)$ is the minimal polynomial of A, then by the above theorem, $m(x)$ must be divisor of $f(x)$.

Cor. 2 : Every root of the minimal equation of a matrix is also a characteristic root of the matrix.

Let $f(x)$ be the characteristic polynomial of a matrix A and $m(x)$ be its minimal polynomial. Then $m(x)$ is a divisor of $f(x)$. Therefore, there exists a polynomial $q(x)$ such that

$$f(x) = m(x)\, q(x) \quad \ldots (1)$$

Suppose λ is a root of the equation $m(x) = 0$. Then $m(\lambda) = 0$. Putting $x = \lambda$ on both sides of (1), we get

$$f(\lambda) = m(\lambda\, q(\lambda) = 0 \cdot q(\lambda) = 0$$

Therefore, λ is also a root of $f(x) = 0$. Thus λ is also a characteristic root of A.

For example, find the minimal polynomial $m(x)$ of matrix A

$$A = \begin{bmatrix} 2 & 2 & -5 \\ 3 & 7 & -15 \\ 1 & 2 & -4 \end{bmatrix}$$

$\Rightarrow$ first find the characteristic polynomial $\Delta(x)$ of A we have

$\text{tr}(A) = 5, \quad A_{11} + A_{22} + A_{33} = 2 - 3 + 8 = 7$

$|A| = 3$

$\therefore \quad \Delta(x) = x^3 - 5x^2 + 7x - 3 = (x-1)^2 (x-3)$

The minimal polynomial m(x) must divides $\Delta(x)$. Also irreducible factor of $\Delta(x)$ are $(x-1)(x-3)$. Also be a factor of m(x).

Thus m(x) is exactly one of the following :

$$f(x) = (x-3)(x-1)$$

or $$g(x) = (x-3)(x-1)^2$$

We know, by Cayley-Hamilton theorem.

$$\Delta(A) = 0$$

$\therefore$ For f(x), $f(A) = (A - I)(A - 3I) = \begin{bmatrix} 1 & 2 & -5 \\ 3 & 6 & -15 \\ 1 & 2 & -5 \end{bmatrix} \begin{bmatrix} -1 & 2 & -5 \\ 3 & 4 & -15 \\ 1 & 2 & -7 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Thus $f(x) = m(x) = (x-1)(x-3) = x^2 - 4x + 3$ is the minimal polynomial of matrix A.

Example :

(1) Consider the following two n-square matrix where $a \neq 0$.

$$J(\lambda, n) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}_{n \times n}$$

and

$$A = \begin{bmatrix} \lambda & a & 0 & \dots & 0 & 0 \\ 0 & \lambda & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & a \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}_{n \times n}$$

The first matrix, called a Jordan block matrix has λ's on the diagonal, 1's are occur above the diagonal entries (superdiagonal element) and 0's elsewhere.

The 2^{nd} matrix A has λ's on the diagonal, as are occur above the diagonal entries and 0's elsewhere then matrix A is generalization of matrix $J(\lambda, n)$ then characteristic and minimal polynomial of matrix $J(\lambda, n)$ and A are same and is equal to $f(x) = (x - \lambda)^n$.

2. Consider a monic polynomial

$$f(x) = x^n + a_{n-1} x^{n-1} + \dots + a_1 t + a_0$$

Let c(f) be a n-square matrix with 1's occurs below the diagonal elements (sub-diagonal) and the negatives of the coefficients in the last column and 0's occur elsewhere.

$$c(f) = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{bmatrix}$$

the c(f) is called comparison matrix of polynomial f(x).

Note : The characteristic polynomial ($\Delta(x)$) and minimal polynomial (m(x)) of comparison matrix are equal to original polynomial f(x).

4.7 Derogatory and Non-derogatory Matrices

An n-rowed matrix is said to be derogatory or non-derogatory, according as the degree of its minimal equation is less than or equal to n.

Thus a matrix is non-derogatory if the degree of its minimal polynomial is equal to the degree of its a characteristic polynomial.

Theorem 19 : Every root of the characteristic equation of a matrix is a root of the minimal equation of the matrix.

Let m(x) be the minimal polynomial of a matrix A. Then m(A) = 0. We know that if λ is a characteristic root of a matrix A and g(x) is any polynomial, then $g(\lambda)$ is a characteristic root of the matrix g(A).

Now, suppose λ is a characteristic root of A. Taking m(x) in place of g(x) we see that $m(\lambda)$ is a characteristic root of m(A). But m(A) is a null matrix and each characteristic root of a null matrix is zero. Therefore, $m(\lambda) = 0 \Rightarrow \lambda$ is a root of the equation m(x) = 0. Hence, every characteristic root of a matrix A is not a root of the minimal polynomial of that matrix.

Theorem 20 : If the roots of the characteristic equation of a matrix are distinct, then the matrix is non-derogatory.

Let A be a matrix of order n whose n characteristic roots are all distinct. As each characteristic root of A is also a root of the minimal polynomial of A, hence in this case the minimal polynomial of A will be of degree n. Consequently, the matrix will be a non-derogatory matrix. In this case, the characteristic equation of A will also give us the minimal equation of A provided we make its leading coefficient unity.

Theorem 21 : The minimal polynomial of an $n \times n$ matrix A is $(-1)^n |A - \lambda I|/g(\lambda)$, where the monic polynomial $g(\lambda)$ is the H.C.F. of the minors of order n – I in $|A - \lambda I|$.

We know that $|A - \lambda I|$ can be expressed as a linear combination of the minors of its any row. Since $g(\lambda)$ is a divisor of every minor of order (n – 1) of $|A - \lambda I|$, therefore, it is also a divisor of $|A - \lambda I|$. So, let

$$|A - \lambda I| = (-1)^n g(\lambda\, h(\lambda) \qquad \dots (1)$$

where $h(\lambda)$ is some monic polynomial. We claim that $h(\lambda)$ is the minimal polynomial of A.

Each element of Adj. $(A - \lambda I)$ is numerically equal to some minor of order $n - 1$ of $|A - \lambda I|$. Let $B(\lambda)$ be the matrix obtained on dividing each element of Adj. $(A - \lambda I)$ by $g(\lambda)$. Then

$$\text{Adj. } (A - \lambda I) = g(\lambda)\, B(\lambda) \quad \ldots (2)$$

Here $B(\lambda)$ is an $n \times n$ matrix and is such that its elements are polynomials in λ having no factor (other than constant) in common.

Pre-multiplying both sides of (2) by $A - \lambda I$, we get

$$(A - \lambda I) \text{ Adj. } (A - \lambda I) = g(\lambda)\, (A - \lambda I)\, B(\lambda) \quad \ldots (3)$$

Since $(A - \lambda I)$ Adj. $(A - \lambda I) = |A - \lambda I|\, I$, we get from (3)

$$|A - \lambda I|\, I = g(\lambda)\, (A - \lambda I)\, B(\lambda)$$

or $$(-1)^n\, g(\lambda)\, h(\lambda)\, I = g(\lambda)\, (A - \lambda I)\, B(\lambda) \quad [\text{from (1)}]$$

Since $g(\lambda) \neq 0$, therefore

$$(-1)^n\, h(\lambda)\, I = (A - \lambda I)\, B(\lambda) \quad \ldots (4)$$

Putting $\lambda = A$ on both sides of (4), we see that $h(A) = 0$. Thus the polynomial $h(\lambda)$ annihilates A.

Let $m(\lambda)$ be the minimal polynomial of A. Then

$$m(A) = 0$$

$$\Rightarrow \quad m(A) - m(\lambda)\, I = -\, m(\lambda)\, I$$

$$\Rightarrow \quad m(A) - m(\lambda I) = -\, m(\lambda)\, I$$

$$\Rightarrow \quad (A - \lambda I)\, L(\lambda) = -\, m(\lambda)\, I$$

where $L(\lambda)$ is a matrix polynomial.

Pre-multiplying both sides of (5) by Adj. $(A - \lambda I)$, we get

$$\Rightarrow \quad \text{Adj. } (A - \lambda I)\, (A - \lambda I)\, L(\lambda) = -\, m(\lambda) \cdot \text{adj. } (A - \lambda I)$$

$$|A - \lambda I|\, I\, L(\lambda) = -\, m(\lambda) \text{ Adj. } (A - \lambda I)$$

or $$(-1)^n\, g(\lambda)\, h(\lambda)\, L(\lambda) = -\, m(\lambda)\, g(\lambda)\, B(\lambda) \quad [\text{from (1) and (2)}]$$

or $$(-1)^n\, h(\lambda)\, L(\lambda) = -\, m(\lambda)\, B(\lambda) \quad [\because\ g(\lambda) \neq 0] \quad \ldots (6)$$

From (6), we see that $h(\lambda)$ is a factor of each element of the matrix $m(\lambda)\, B(\lambda)$. But the elements of $B(\lambda)$ have no factor in common. Therefore, $h(\lambda)$ must be a divisor of $m(\lambda)$. But $h(\lambda)$ annihilates and $m(\lambda)$ is the minimal polynomial of A. Therefore, $m(\lambda)$ is also a divisor of $h(\lambda)$. Since both $h(\lambda)$ and $m(\lambda)$ are monic polynomials, therefore must have $m(\lambda) = h(\lambda)$.

SOLVED EXAMPLE

Example 4.28 : Show that the matrix $A = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}$ is derogatory.

Solution : We have,

$$|A - \lambda I| = \begin{bmatrix} 7 - \lambda & 4 & -1 \\ 4 & 7 - \lambda & -1 \\ -4 & -4 & 4 - \lambda \end{bmatrix} = -\,(3 - \lambda)^2\, (\lambda - 12)$$

Therefore, the characteristic roots of A are 3, 3, 12.

We know that each characteristic root of A is not a root of its minimal polynomial. So if $m(\lambda)$ is the minimal polynomial of, then both $(x - 3)$ and $(x - 12)$ are factors of $m(\lambda)$. Let us try whether the polynomial $h(x) = (x - 3)(x - 12) = x^2 - 15x + 36$ annihilates or not.

We have, $$A^2 = \begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix}$$

$$\therefore \quad A^2 - 15A + 36I = \begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix} - \begin{bmatrix} 105 & 60 & -15 \\ 60 & 105 & -15 \\ -60 & -60 & 60 \end{bmatrix} + \begin{bmatrix} 36 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 36 \end{bmatrix}$$

$$= 0$$

$\therefore$ $h(x)$ annihilates A. Thus $h(x)$ is the monic polynomial of lowest degree which annihilates A. Since its degree is less than the order of A, hence A is derogatory.

Example 4.29 : Show that every unit matrix of order ≥ 2 is derogatory.

Solution : Let I be a unit matrix of order n where $n \geq 2$. We see that the polynomial $m(x) = x - 1$ annihilates I. Therefore, $x - 1$ is the minimal polynomial of I. Since degree of $x - 1$ is 1 which is less than n, therefore I is derogatory.

Example 4.30 : Find the minimal polynomial of the $n \times n$ matrix A each of whose elements is I.

Solution : Let A be the $n \times n$ matrix each of whose elements is 1. Then $A^2 = nA$. Therefore, the polynomial $x^2 - nx$ annihilates A. Now the polynomial $x + a$ cannot annihilates A whatever may be the value of the scalar a. Therefore, $x^2 - nx$ is the monic polynomial of lowest degree which annihilates A. Hence $x^2 - nx$ is the minimal polynomial of A.

Example 4.31 : If A and B be $n \times n$ matrices and B be non-singular, then show that A and $B^{-1}AB$ have the same minimal polynomial.

Solution : First we shall show that a monic polynomial $f(x)$ annihilates A if, and only if, it annihilates $B^{-1}AB$. We have

$$(B^{-1}AB)^2 = B^{-1}ABB^{-1}AB = B^{-1}A^2B$$

Proceeding in this way we can show that $(B^{-1}AB)^k = B^{-1}A^kB$, where k is any positive integer.

Let $f(x) = x^r + a_1x^{r-1} + \ldots + a_{r-1}x + a_r$

Then $f(A) = A^r + a_1A^{r-1} + \ldots + a_{r-1}A + a_rI$

Also, $f(B^{-1}AB) = (B^{-1}AB)' + a_1 (B^{-1}AB)^{r-1} + \ldots + a_{r-1} (B^{-1}AB) + aI$

$= B^{-1}A^rB + a_1B^{-1} + A^{r-1}B + \ldots + a_{r-1} (B^{-1}AB) + a_rB^{-1}B$

$= B^{-1}(A^r + a_1A^{r-1} + \ldots + a_{r-1}A + a_rI) B$

$= B^{-1} f(A) B$

Since B is non-singular, therefore

$B^{-1} f(A) B = 0$ if, and only if $f(A) = 0$.

Thus f(x) annihilates A if, and only if, it annihilates $B^{-1}AB$.

Therefore if f(x) is the polynomial of lowest degree that annihilates A then it is also the polynomial of lowest degree that annihilates $B^{-1}AB$ and conversely.

Hence A and $B^{-1}AB$ have the same minimal polynomial.

Example 4.32 : A square matrix is said to be idempotent if $A^2 = A$. Show that if A is idempotent, then all eigenvalues of A are equal to 1 or 0.

Solution : We have, $A^2 = A \Rightarrow A^2 - A = 0$

$\Rightarrow$ A satisfies the polynomial $\lambda^2 - \lambda$

$\Rightarrow$ $\lambda^2 - \lambda$ annihilates A.

$\Rightarrow$ the minimal polynomial, $m(\lambda)$ of A divides $\lambda(\lambda - 1)$.

$\Rightarrow$ $m(\lambda) = \lambda, \lambda - 1$ or $\lambda(\lambda - 1)$.

$\Rightarrow$ $\lambda = 0$ or 1 are the only roots of $m(\lambda)$.

As each eigenvalue of A is also a root of its minimal polynomial. Hence, eigenvalues of A are equal to 1 or 0.

Remark : Null matrix is the only matrix whose minimal polynomial is Unit matrix is the only matrix whose minimal polynomial is $\lambda - 1$. Therefore, if $A^2 = A$ and $A \neq 0$ and $A \neq I$, then the minimal polynomial of A is $\lambda^2 - \lambda$.

Exercise 4.3

Q. (I)

1. Find the minimal polynomial of the matrix $\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$ and show that it is derogatory.

2. Show that the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$ is derogatory. Thos that the minimal polynomial of a non-zero idempotent matrix C ($\neq I$) is $\lambda^2 - \lambda$.

3. Determine the minimal and characteristic equations of the following matrices :

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

Show that the matrix $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$ is non-derogatory.

Q. II

1. Find the characteristic polynomials of the following matrices :

(a) $\begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}$

(e) $\begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} -2 & -7 \\ 1 & 2 \end{bmatrix}$ (g) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

2. Find the eigenvalues of the matrices in exercise 1.
3. Find the eigenspaces and bases for each eigenvalue of the matrices in exercise 1.
4. Find the characteristic polynomials of the following matrices :

(a) $\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -7 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{bmatrix}$ (c) $\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$

(d) $\begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} -2 & 0 & 1 \\ -6 & -2 & 0 \\ 19 & 5 & -4 \end{bmatrix}$ (f) $\begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix}$

(g) $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 3 & 0 \\ -4 & 13 & -1 \end{bmatrix}$ (h) $\begin{bmatrix} 5 & 0 & 1 \\ 1 & 1 & 0 \\ -7 & 1 & 0 \end{bmatrix}$

5. Find the eigenvalues of the matrices in exercise 4.
6. Find the eigenspaces and bases for each eigenvalue of the matrices in exercise 4.
7. Find the eigenvalues and bases for the eigenspaces of

(a) A^{25}, where,

$$A = \begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

(b) A^7, where $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$

8. Prove that $\lambda = 0$ is an eigenvalue of a matrix A if and only if A is not invertible.

9. Prove that the constant term in the characteristic polynomial of an $n \times n$ matrix A is $(-1)^n \det(A)$.
10. Prove that a square matrix A and its transpose A^t have the same characteristic polynomial.
11. If A and B are $n \times n$ matrices, with A non-singular, prove that AB and BA have the same set of eigenvalues.
12. Show that the following matrices are not diagonalizable :

(a) $\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -3 \\ 1 & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$

(d) $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$ (e) $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 3 & 0 \\ -4 & 13 & -1 \end{bmatrix}$.

In Exercises 13 – 18, find a matrix P that diagonalizes A, and determine $P^{-1}AP$.

13. $A = \begin{bmatrix} -14 & 12 \\ -20 & 17 \end{bmatrix}$ 14. $A = \begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$ 15. $\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$

16. $A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$ 17. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ 18. $A = \begin{bmatrix} 2 & 0 & -2 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

In Exercises 19 - 24, determine whether A is diagonalizable. If so, find a matrix P that diagonalizes A, and determine $P^{-1}AP$.

19. $A = \begin{bmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{bmatrix}$ 20. $A = \begin{bmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{bmatrix}$

21. $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$ 22. $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}$

23. $A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}$ 24. $A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$

25. Find det (A), given that $P(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4$ is the characteristics of A.
26. Find the geometric and algebraic multiplicity of the matrices in exercise 20, 22, 24.
27. Find A^{14}, where $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$.

28. Find A^{11}, where $A = \begin{bmatrix} -1 & 7 & -1 \\ 0 & 1 & 0 \\ 0 & 15 & -2 \end{bmatrix}$. 29. Compute A^{10}, where $A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$.

Answers (4.3)

1. (a) $\lambda^2 - 3\lambda + 2$ (b) $\lambda^2 - 2\lambda + 1$ (c) $\lambda^2 - 2\lambda$
   (d) $\lambda^2 - 8\lambda + 16$ (e) $\lambda^2 - 12$ (f) $\lambda^2 + 3$
   (g) $\lambda^2 - 2\lambda + 1$
2. (a) $\lambda = 1, \lambda = 2$ (b) $\lambda = 1$ (c) $\lambda = 0, \lambda = 2$.
   (d) $\lambda = 4$ (e) $\lambda = 2\sqrt{3}$, $\lambda = -2\sqrt{3}$ (f) No real eigenvalues.
   (g) $\lambda = 1$.
3. (a) $E(\lambda = 1) = \{t(-1, 1) \mid t \in \mathbb{R}\}$, basis is $\{(-1, 1)\}$.
   $E(\lambda = 2) = \{t(-2, 1) \mid t \in \mathbb{R}\}$, basis is $\{(-2, 1)\}$.
   (b) $E(\lambda = 1) = \{t(1, 0) \mid t \in \mathbb{R}\}$, basis is $\{(-1, 0)\}$.
   (c) $E(\lambda = 2) = \{t(1, 1) \mid t \in \mathbb{R}\}$, basis is $\{(1, 1)\}$.
   $E(\lambda = 0) = \{t(1, -1) \mid t \in \mathbb{R}\}$, basis is $\{(1, -1)\}$
   (d) $E(\lambda = 4) = \{t(3/2, 1) \mid t \in \mathbb{R}\}$, basis is $\{(3/2, 1)\}$.
   (e) $E(\lambda = 2\sqrt{3}) = \{t(3/\sqrt{12}, 1) \mid t \in \mathbb{R}\}$, basis is $\{(3/\sqrt{12}, 1)\}$.
   $E(\lambda = -\sqrt{3}) = \{t(-3/\sqrt{12}, 1)\} : \{(-3/\sqrt{12}, 1)\}$.
   (f) There are no eigenspaces.
   (g) $E(\lambda = 1) = \{ r(1, 0) + s(0, 1) \mid r, s \in \mathbb{R}\} ; \{(1, 0), (0, 1)\}$.
4. (a) $(\lambda - 1)^3$ (b) $(\lambda - 1)(\lambda - 2)(\lambda - 21)$
   (c) $(\lambda - 1)(\lambda - 2)^2$ (d) $\lambda^3 - 6\lambda^2 + 11\lambda - 6$
   (e) $\lambda^3 + 8\lambda^2 + \lambda + 8$ (f) $\lambda^3 - 2\lambda^2 - 15\lambda + 36$
   (g) $\lambda^3 - \lambda^2 - \lambda - 2$ (h) $\lambda^3 - 6\lambda^2 + 12\lambda - 8$
5. (a) $\lambda = 1, \lambda = 1, \lambda = 1$ (b) $\lambda = 1, \lambda = 2, \lambda = 21$
   (c) $\lambda = 1, \lambda = 2, \lambda = 2$. (d) $\lambda = 1, \lambda = 2, \lambda = 3$
   (e) $\lambda = -8$ (f) $\lambda = -4, \lambda = 3$
   (g) $\lambda = 2$ (h) $\lambda = 2$.
6. (a) $E(\lambda = 1) = \{t(0, 0, 1) \mid t \in \mathbb{R}\} ; \{(0, 0, 1)\}$
   (b) $E(\lambda = 1) = \{ t(1, -1, 0) \mid t \in \mathbb{R}\} ; \{(1, -1, 0)\}$
   $E(\lambda = 2) = \{t(3, 3, -1) \mid t \in \mathbb{R}\} ; \{(3, 3, -1)\}$
   $E(\lambda = 21) = \{t(1, 1, 6) \mid t \in \mathbb{R}\} ; \{(1, 1, 6)\}$.
   (c) $E(\lambda = 1) = \{t(3, -1, 3) \mid t \in \mathbb{R}\} ; \{(3, -1, 3)\}$
   $E(\lambda = 2) = \{t(2, 1, -1) \mid t \in \mathbb{R}\} ; \{(2, 2, -1)\}$.
   (d) $E(\lambda = 1) = \{t(0, 1, 0) \mid t \in \mathbb{R}\} ; \{(0, 1, 0)\}$
   $E(\lambda = 2) = \{t(-1/2, 1, 1) \mid t \in \mathbb{R}\} ; \{(-1/2, 1, 1)\}$.

$E(\lambda = 3) = \{t(-1, 1, 1) \mid t \in \mathbb{R}\}$; $\{(-1, 1, 1)\}$.

(e) $E(\lambda = -8) = \{t(-1/6, -1/6, 1) \mid t \in \mathbb{R}\}$; $\{(-1/6, -1/6, 1)\}$

(f) $E(\lambda = -4) = \{t(-2, 8/3, 1) \mid t \in \mathbb{R}\}$; $\{(-2, 8/3, 1)\}$

$E(\lambda = 3) = \{t(5, -2, 1) \mid t \in \mathbb{R}\}$; $\{(5, -2, 1)\}$

(g) $E(\lambda = 2) = \{t(1/3, 1/3, 1) \mid t \in \mathbb{R}\}$; $\{(1/3, 1/3, 1)\}$

(h) $E(\lambda = 2) = \{t(-1/3, -1/3, 1) \mid t \in \mathbb{R}\}$; $\{(-1/3, -1/3, 1)\}$.

7. (a) $\lambda = 1, \lambda = -2$;

basis for $\lambda = 1$ is $\{(-1, 1, 0), (-1, 0, 1)\}$

basis for $\lambda = -1$ is $\{(2, -1, 1)\}$.

(b) $\lambda = 128, \lambda = 1$;

basis for $\lambda = 128$ is $\{(-1, 0, 1), (0, 1, 0)\}$

basis for $\lambda = 1$ is $\{(-2, 1, 1)\}$.

13. $P = \begin{bmatrix} \frac{4}{5} & \frac{3}{4} \\ 1 & 1 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

14. $P = \begin{bmatrix} \frac{1}{3} & 0 \\ 1 & 1 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

15. $P = \begin{bmatrix} \frac{1}{2} & 0 \\ 1 & 1 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

16. $P = \begin{bmatrix} -1 & 0 & -2 \\ 1 & -1 & -3 \\ 0 & 1 & 1 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

17. $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

18. $P = \begin{bmatrix} -2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

19. Not diagonalizable.

20. $P = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

21. Not diagonalizable.

22. $P = \begin{bmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

23. Not diagonalizable.

24. $P = \begin{bmatrix} -1 & -1 & \frac{1}{3} \\ 1 & 0 & -\frac{2}{3} \\ 0 & 1 & 0 \end{bmatrix}$; $P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}$.

25. 4

26. Geometric multiplicity algebraic multiplicity.

(20)	$\lambda = 1$	1	1
	$\lambda = 2$	1	1
	$\lambda = 3$	1	1
(22)	$\lambda = 0$	2	2
	$\lambda = 1$	1	1
(24)	$\lambda = 1$	2	2
	$\lambda = 7$	1	1

27. $A^{13} = \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix}$

28. $A^{11} = \begin{bmatrix} -1 & 10237 & -2047 \\ 0 & 1 & 0 \\ 0 & 10245 & -2048 \end{bmatrix}$

29. $A^{10} = \begin{bmatrix} 1 & 0 \\ 1023 & 1024 \end{bmatrix}$

Important Points

- Eigenvalues and eigenvectors.
- Equivalent statements relating an eigenvalue of a matrix.
- Eigenspaces corresponding to eigenvalues.
- Finding, eigenvalues, corresponding eigenspaces, bases and dimensions of eigenspace.
- Eigenvalues of power of a matrix.
- Eigenvalues and invertibility of a matrix.
- Diagonalization of a matrix.
- Necessary and sufficient condition for a matrix to be diagonalizable.

- Procedure to determine the given matrix is diagonalizable, finding matrix P that diagonalises the matrix.
- Geometric and algebraic multiplicities of eigenvalues.
- An application of diagonalization to find higher power of a matrix.

Miscellaneous Exercise

(A) Theory Questions :

1. If A is $n \times n$ triangular matrix, then prove that the eigenvalues of A are the entries on the main diagonal of A.
2. Let A be an $n \times n$ matrix and λ be a real number.
   (a) Prove that λ is an eigenvalue of A if and only if the system $(\lambda I - A) X = 0$ has non-trivial solutions.
   (b) Prove that there is a non-zero vector $\vec{x}$ in $\mathbb{R}^n$ such that $A\vec{x} = \lambda\vec{x}$ if and only if λ is a solution of the equation $\det(\lambda I - A) = 0$.
3. If k is positive integer, λ is an eigenvalue of a matrix A and $\vec{x}$ is a corresponding eigenvector, then prove that λ^k is an eigenvalue of A^k and $\vec{x}$ is a corresponding eigenvector.
4. A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.
5. Prove that the constant term in the characteristic polynomial of an $n \times n$ matrix is $(-1)^n \det A$.
6. The characteristic roots of a diagonal matrix are the same as its diagonal elements.
7. If $A = \begin{bmatrix} P & Q \\ O & R \end{bmatrix}$ where A, P are square matrices, show that the characteristic roots of A are the characteristic roots of P and R taken together and conversely.
8. Show that at least one characteristic root of every singular matrix is 0.
9. If λ is a characteristic root of a non-singular matrix A, then λ^{-1} is a characteristic of A^{-1}.

(B) Numerical Problems :

1. Find the eigenvalues and eigenvectors of the following matrices. In each case compute the eigenvectors to find matrix P and verify that $P^{-1}AP = D$ where D is diagonal matrix of the eigenvalues.

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & -2 & 0 \\ -2 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 4 \\ 2 & -4 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 9 & 15 & 3 \\ 6 & 10 & 2 \\ 3 & 5 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}.$$

2. Let matrix A and B are $A = \begin{bmatrix} 1 & 4 & 1 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 9 & 15 & 3 \\ 6 & 10 & 2 \\ 3 & 5 & 1 \end{bmatrix}$.
   Show that $B = A^2 + A$ and if λ is eigenvalue of A then $(\lambda^2 + \lambda)$ is a eigen value of B.
3. Using the characteristic equation, show taht λ is an eigenvalue of matrix A then $\frac{1}{1+\lambda}$ is an eigenvalues of $\frac{1}{(I+A)}$.
4. If eigenvalues of matrix A are positive then prove that eigenvalues of matrix $(A + A^{-1})$ are equal or greater than 2.
5. Prove that sum of eigenvalues of a matrix equal to its trace and their product equal to it's determinant.
6. Prove that eigenvalues of real symmetric matrix are real.
7. Prove that eigenvectors of a symmetric matrix X are orthogonal to each other.
8. Suppose A is idempotent and symmetric. Prove the following :
   (a) eigenvalues are 0 or 1. (b) trace equal to its rank.
   (c) A = BB' where B'B = I.
9. Prove that eigenvalues of BA are the same as those of ABA when A is idempotent.
10. Prove that the only non-singular idempotent matrix is the identity matrix.
11. For any idempotent matrix A, prove that
    (a) A^k has the same eigenvalues as A. (b) A^k has rank r_A and (r_A = rank (A)).
    (c) Rank $(I - A) = n - r_A$, for A of order n.

❐❐❐

Chapter 5...

Linear Transformations

Synopsis

5.1 Introduction
5.2 The Definition of Linear Transformation and Examples
5.3 Sum Scalar Multiple and Composition of Linear Transformation
5.4 Kernel and Range
5.5 Inverse of a Linear Transformation
5.6 Matrix of a Linear Transformation

Objectives

At the end of this chapter.

(1) You will be able to understand the linear transformation, inverse linear transformation also. Kernel and range of linear transformation.

Gottfried Wilhelm Leibniz

Gottfried Wilhelm (von) Leibniz French: Godefroi Guillaume Leibnitz; 1 July 1646 [O.S. 21 June] – 14 November 1716) was a German polymath and philosopher who occupies a prominent place in the history of mathematics and the history of philosophy, having developed differential and integral calculus independently of Isaac Newton. Leibniz's notation has been widely used ever since it was published. It was only in the 20th century that his Law of Continuity and Transcendental Law of Homogeneity found mathematical implementation (by means of non-standard analysis). He became one of the most prolific inventors in the field of mechanical calculators. While working on adding automatic multiplication and division to Pascal's calculator, he was the first to describe a pinwheel calculator in 1685[13] and invented the Leibniz wheel, used in the arithmometer, the first mass-produced mechanical calculator. He also refined the binary number system, which is the foundation of virtually all digital computers.

5.1 Introduction

In this chapter we shall study functions from one vector space to another which preserve the addition and scalar multiplication, that is addition of two vectors in domain vector space map to sum of their respective images in codomain vector space, the scalar multiple of a vector in domain vector space maps to scalar multiple of the image of the vector in codomain vector space. All this is generalization of addition of matrices and scalar multiplication of matrices. We will see the linear transformations which are

(i) **one-to-one or injective** (ii) **onto** or **surjective** and (iii) **one-to-one** and **onto** or **bijective**. We will see the relations corresponding to the matrices, left inverse, right inverse or inverse of a matrix.

Overall, we will try to establish that there is analogy between the matrices and linear transformations. This will be done by defining matrix representation of a linear transformation from a finite dimensional vector space to a finite dimensional vector space.

Note that some authors use V_n and W_m to denote vector spaces of dimension n and m respectively, but we will use V and W simply and will mention dimensions whenever necessary. Another one more thing that some authors write vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ as vertical columns matrices or the transpose of the row matrix. For example, a vector (1, 3, 2) (we use this notation) in $\mathbb{R}^n$ or $\mathbb{C}^n$ is written as $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ or $[1\ \ 3\ \ 2]^T$ by some authors. However, we will be using ordered triples or n-tuples to denote vectors in $\mathbb{R}^3$ or $\mathbb{C}^n$ for simplicity and for use of minimum space.

5.2 The Definition of Linear Transformation and Examples

Definition : Let V and W be two vectors spaces. A function $T : V \to W$ from V to W is said to be a linear transformation if

(a) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$, for all $\vec{u}$ and $\vec{v}$ in V.

and (b) $T(k\vec{u}) = kT(\vec{u})$, for all $\vec{u} \in V$ and for all scalars k.

SOLVED EXAMPLES

Example 5.1 : Let F be a function defined by the formula $F(x, y) = (2x, x + y, x - y)$ from $\mathbb{R}^2$ into R^3. Then F is a linear transformation.

Solution : If $\vec{u} = (x_1, y_1)$ and $\vec{v} = (x_2, y_2)$

then $\vec{u} + \vec{v} = (x_1 + x_2, y_1 + y_2)$

so that $F(\vec{u} + \vec{v}) = F(x_1 + x_2, y_1 + y_2)$

$= (2(x_1 + x_2), (x_1 + x_2) + (y_1 + y_2), (x_1 + x_2) - (y_1 + y_2))$

[We let $x = x_1 + x_2$ and $y = y_1 + y_2$ in the given formula]

$= (2x_1 + 2x_2, (x_1 + y_1) + (x_2 + y_2), (x_1 - y_1) + (x_2 - y_2))$

$= (2x_1, x_1 + y_1, x_1 - y_1) + (2x_2, x_2 + y_2, x_2 - y_2)$

$= F(x_1, y_1) + F(x_2, y_2)$

$= F(\vec{u}) + F(\vec{v})$

Moreover, if k is any scalar, then $k(x_1, y_1) = (kx_1 , ky_1)$, so that

$$F(k\vec{u}) = F(kx_1 , ky_1)$$

$$= (2kx_1 , kx_1 + ky_1 , kx_1 - ky_1)$$ [We let $x = kx_1$ and $y = ky_1$ in the given formula]

$$= (k(2x_1), k(x_1 + y_1), k(x_1 - y_1))$$
$$= k\,(2x_1, x_1 + y_1 , x_1 - y_1)$$

$$= kF\,(x_1 , y_1) = kF(\vec{u})$$

Thus, F is a linear transformation.

Let us prove the following result.

Result 1 : If $F : V \to W$ is a function, then F is linear transformation if and only if for any vectors $\vec{u_1}$ and $\vec{u_2}$ in V and any scalars k_1 and k_2

$$F(k_1\vec{u_1} + k_2\vec{u_2}) = k_1F(\vec{u_1}) + k_2F(\vec{u_2})$$

Proof : Suppose $F : V \to W$ is a linear transformation. Then, for any $\vec{u_1}$ and $\vec{u_2}$ in V and any scalars k_1 and k_2, we have

$$F\,(k_1\vec{u_1} + k_2\vec{u_2}) = F(k_1\vec{u_1}) + F(k_2\,\vec{u_2}), \quad \text{by (a) of the definition}$$

$$= k_1F(\vec{u_1}) + k_2\,F(\vec{u_2}), \quad \text{by (b) of the definition}$$

Thus, the condition follows.

Conversely, suppose for any $\vec{u_1}$ and $\vec{u_2}$ in V and any scalars k_1 and k_2, we have

$$F(k_1\vec{u_1} + k_2\vec{u_2}) = k_1F(\vec{u_1}) + k_2F\,(\vec{u_2}) \quad \ldots (*)$$

To show that F is linear transformation.

Since (*) holds for all scalars k_1 and k_2, let $k_1 = k_2 = 1$ in (*), we get

$$F(1.\vec{u_1} + 1.\vec{u_2}) = 1.F(\vec{u_1}) + 1.F(\vec{u_2})$$

that is, $F(\vec{u_1} + \vec{u_2}) = F(\vec{u_1}) + F(\vec{u_2})$

Also, taking $k_1 = k$ and $k_2 = 0$ in (*) we get $F(k\vec{u_1}) = kF(\vec{u_1})$. Thus, the conditions (a) and (b) of the definition of linear transformation hold, therefore F is a linear transformation.

Note : 1. From result 1, we can define linear transformation equivalently as : "A function $F : V \to W$ is a linear transformation if for all $\vec{u_1}$ and $\vec{u_2}$ in V and any scalars k_1 and k_2

$$F(k_1\vec{u_1} + k_2\vec{u_2}) = k_1F(\vec{u_1}) + k_2F(\vec{u_2}). \quad \ldots (*)$$

2. If $F : V \to W$ is linear transformation, the identity (*) can be generalised as : For any vectors $\overrightarrow{u_1}$, $\overrightarrow{u_2}$, … $\overrightarrow{u_n}$ in V and any scalars k_1 , k_2 , … , k_n,

$$F(k_1\overrightarrow{u_1} + k_2\overrightarrow{u_2} + \ldots + k_n\overrightarrow{u_n}) = k_1F(\overrightarrow{u_1}) + k_2F(\overrightarrow{u_2}) + \ldots + k_n F(\overrightarrow{u_n}).$$

3. If F_1 and F_2 are two linear transformations from V into W, then F_1 and F_2 will be called equal, written as $F_1 = F_2$, if

$$F_1(\overrightarrow{u}) = F_2(\overrightarrow{u}) \text{ for every } \overrightarrow{u} \text{ in V.}$$

Example 5.2 : If A is a fixed $m \times n$ matrix say

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ : & : & & : \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{bmatrix} \text{ and } X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

is any vector in R^n then AX is a vector in R^m, and the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by the formula :

$$T_A(X) = AX = \begin{bmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ : & : & & : \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix} \quad \ldots (1)$$

is a linear transformation called *multiplication by A.* Such linear transformations are called *matrix transformations.*

Solution : Let u and v be any $n \times 1$ matrices and k be any scalar. Using the properties of matrix multiplication, we have

$$T_A(u + v) = A(u + v) = Au + Av = T_A(u) + T_A(v)$$

and $$T_A(ku) = A(ku) = k(Au) = kT_A(u)$$

which proves that T_A is a linear transformation.

Remark : The notations $X = (x_1 , x_2 , \ldots , x_n)$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$ will both be used for vectors in $\mathbb{R}^n$ or $\mathbb{C}^n$ according to the convenience.

Example 5.3 : Let θ be a fixed angle, and let $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ be multiplication by the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ +\sin\theta & \cos\theta \end{bmatrix}$$

i.e. if $\vec{u}$ is the vector $\vec{u} = \begin{bmatrix} x \\ y \end{bmatrix}$

then $T_A(\vec{u}) = A\vec{u} = \begin{bmatrix} \cos\theta & -\sin\theta \\ +\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}$

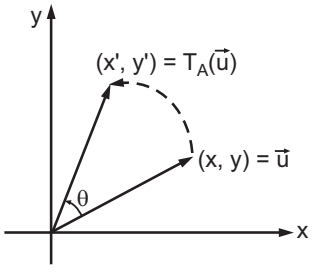


Fig. 5.1

Geometrically, $T_A(\vec{u})$ is the vector that results if $\vec{u}$ is rotated counterclockwise through an angle θ about origin.

Solution : To show that T_A is linear transformation (note that T_A is a particular case of example (2) is left as an exercise for the students).

Example 5.4 : (Zero transformation) : Let V and W be any two vector spaces. The mapping $T : V \to W$ given by $T(\vec{u}) = \vec{0}$ for every $\vec{u}$ in V is a linear transformation.

Solution : For any $\vec{u}$ and $\vec{v}$ in V,

$$T(\vec{u} + \vec{v}) = \vec{0} = \vec{0} + \vec{0}$$

$$= T(\vec{u}) + T(\vec{v}), \text{ since } T(\vec{u}) = \vec{0} \text{ and } T(\vec{v}) = \vec{0}$$

Also, for any scalar k,

$$T(k\vec{u}) = \vec{0} = k \cdot \vec{0} = k.T(\vec{u}) \text{ (as } T(\vec{u}) = \vec{0})$$

Example 5.5 : (Identity transformation) : For any vector space V, the mapping $I : V \to V$ defined by $I(\vec{u}) = \vec{u}$ is called the *identity transformation* on V. It is easy to check that I is a linear transformation. For if $\vec{u}$ and $\vec{v}$ are any vectors in V, then

$$I(\vec{u} + \vec{v}) = \vec{u} + \vec{v} = I(\vec{u}) + I(\vec{v})$$

Also, for any scalar k,

$$I(k\vec{u}) = k\vec{u} = kI(\vec{u})$$

Example 5.6 : Let V be any vector space and k fixed scalar. A function $T : V \to V$ defined by $T(\vec{u}) = k\vec{u}$, for $\vec{u}$ in V, is a linear transformation on V.

Solution : For any $\vec{u}$ and $\vec{v}$ in V

$$\begin{aligned} T(\vec{u} + \vec{v}) &= k(\vec{u} + \vec{v}) \\ &= k\vec{u} + k\vec{v} \\ &= T(\vec{u}) + T(\vec{v}) \end{aligned}$$

Also, for any scalar α,

$$\begin{aligned} T(\alpha\vec{u}) &= k(\alpha\vec{u}) = (k\alpha)\vec{u} \\ &= (\alpha k)\vec{u} = \alpha(k\vec{u}) \\ &= \alpha T(\vec{u}) \end{aligned}$$

Remark : The above linear transformation is called a dilation of V with factor k if $k > 1$ and is called a contraction of V with factor k if $0 < k < 1$.

Example 5.7 : We look at some particular cases of example (2).

(i) If $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, then the map $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ given by,

$T_A(X) = AX, \quad X \in \mathbb{R}^2, \qquad X$ is considered as 2×1 column matrix.

Then clearly T_A is a linear transformation of $\mathbb{R}^2$ onto $\mathbb{R}^2$. Observe that this linear map is **reflection of** X **with respect to Y-axis.**

(ii) If $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, then map $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T_A(X) = AX$ is a **reflection of** X **with respect to X-axis.**

(iii) If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, then the map $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T_A(X) = AX$ is a **projection of** X **on X-axis.**

(iv) If $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, then the map $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T_A(X) = AX$ is a **projection of** X **on Y-axis.**

(v) If $A = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$, then the map $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is linear map and it is stretching along X-axis if $k > 1$.

(vi) If $A = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$, then the map $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map and it is stretching along Y-axis if $k > 1$.

Example 5.8 : (i) $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by $T(x, y, z) = (x, yz)$ is not a linear map.

(ii) $T : \mathbb{R}^2 \to \mathbb{R}^3$ given by $T(x, y) = (x, x + y, y + 3)$ is not a linear map.

(iii) $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T(x, y, z) = (x^2, z, y)$ is not a linear map.

(iv) $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(x, y) = (|x|, |y|)$ is not a linear map.

(v) $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a map given by $T(x, y, z) = (1, 1, 1)$, then T is not a linear map.

Exercise : Students are requested to prove each part in example (8) with reasons.

Example 5.9 : Let V be a vector space of all functions defined on [0, 1] and W be subspace of V consisting of all continuously differentiable functions on [0, 1]. Let $D : W \to V$ be the map which maps $f = f(x)$ into its derivative; that is,

$$D(f) = f'(x)$$

Then show that D is a linear transformation.

Solution : For f and g in W

$$\begin{aligned} D(f + g) &= (f + g)'(x) \\ &= f'(x) + g'(x) && \text{by property of differentiation} \\ &= D(f) + D(g) \end{aligned}$$

Also, for any scalar k,

$$\begin{aligned} D(kf) &= (kf)'(x) \\ &= kf'(x) && \text{by property of differentiation} \\ &= kD(f) \end{aligned}$$

So that D is a linear transformation.

Example 5.10 : Let $p = p(x) = a_0 + a_1x + \ldots + a_nx^n$ be a polynomial in P_n, and define the map $T : P_n \to P_{n+1}$ by $T(p) = T(p(x)) = xp(x) = a_0x + a_1x^2 + \ldots\ldots + a_n x^{n+1}$. Then show that T is a linear transformation.

Solution : For any two polynomials p_1 and p_2,

$$\begin{aligned} T(p_1 + p_2) &= T\big(p_1(x) + p_2(x)\big) = x\big(p_1(x) + p_2(x)\big) \\ &= xp_1(x) + xp_2(x) \\ &= T(p_1) + T(p_2) \end{aligned}$$

Also, for any scalar k

$$T(kp_1) = T\big(kp_1(x)\big) = x\big(kp_1(x)\big) = k\big(xp_1(x)\big) = kT(p_1)$$

So that T is a linear transformation.

Example 5.11 : Let V be a vector space of continuous functions on [0, 1] and define a map $T : V \to \mathbb{R}$ by, for f in V.

$$T(f) = \int_0^1 f(x)\,dx$$

Show that T is a linear transformation.

Solution : For any f and g in V

$$T(f+g) = \int_0^1 (f+g)(x)\, dx = \int_0^1 (f(x)+g(x))\, dx$$

$$= \int_0^1 f(x)\, dx + \int_0^1 g(x)\, dx$$

$$= T(f) + T(g)$$

Also, for any scalar k,

$$T(kf) = \int_0^1 (kf)(x)\, dx = \int_0^1 k\, f(x)\, dx$$

$$= k \int_0^1 f(x)\, dx = kT(f)$$

Domain, Co-domain, Range of a Linear Transformation :

Definition : Let $T : V \to W$ be a linear transformation and $\vec{u} \in V$ be a vector in V. Then $\vec{w} = T(\vec{u})$ is called the **image of $\vec{u}$** under T. If S is a subspace in V, then we denote by T(S) = The set of all vectors $T(\vec{u})$ in W, where $\vec{u} \in S$. We call T(S)**, the image of** S **in** W **under** T**.**

The vector space V is called **domain of** T and T(V) is called the **range of** T and W is called the **codomain of** T.

Note : In later stage we will prove that the range of T, T(V) is a subspace of W.

Definition : A linear transformation $T : V \to W$ is said to be **onto** (or surjective) if the range of T is the entire vector space W.

Again, if every vector $\vec{v}$ in T(V), which is the range of T lying in W is an image of exactly one element of domain vector space V, then T is called **one-to-one** (injective) linear transformation. In other words, T is one-to-one if distinct vectors of V are transformed into distinct vectors of W. Symbolically this can be stated as if $\vec{u}$ and $\vec{v}$ are in V, then $T(\vec{u}) = T(\vec{v})$ implies that $\vec{u} = \vec{v}$.

A one-to-one and onto linear transformation is called **bijective**. **A linear transformation is also called homomorphism**. A one-to-one and onto linear transformation is called an **isomorphism.**

SOLVED EXAMPLES

Example 5.12 : Let a map $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$T(x_1, x_2, x_3) = (x_1, x_2) \text{ for } (x_1, x_2, x_3) \in \mathbb{R}^3$$

is a linear transformation. This linear transformation is called **projection.**

It is **onto** but **not one-to-one.**

Solution : For any scalars k and l and for any $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$ in $\mathbb{R}^3$,

$$\begin{aligned} k\vec{u} + l\vec{v} &= k(u_1, u_2, u_3) + l(v_1, v_2, v_3) \\ &= (ku_1 + lv_1, ku_2 + lv_2, ku_3 + lv_3) \end{aligned}$$

Therefore, by definition of T, we have

$$\begin{aligned} T(k\vec{u} + l\vec{v}) &= (ku_1 + lv_1, ku_2 + lv_2) \\ &= (ku_1, ku_2) + (lv_1, lv_2) \\ &= k(u_1, u_2) + l(v_1, v_2) \\ &= kT(\vec{u}) + lT(\vec{v}) \end{aligned}$$

Thus T is a linear transformation.

We observe that for any $(u, v) \in \mathbb{R}^2$, we have $(u, v, w) \in \mathbb{R}^3$, whatever $w \in \mathbb{R}$ such that $T(u, v, w) = (u, v)$, hence T is **onto**.

This itself shows that there are more than two vectors (infact infinitely many) in $\mathbb{R}^3$ which map the same vector (u, v) in $\mathbb{R}^2$, hence T is **not one-to-one**.

It is also to be noted that the range of T is whole of $\mathbb{R}^2$, that is $T(\mathbb{R}^3) = \mathbb{R}^2$.

Example 5.13 : Recall the examples (3), (4), (5) and (6), in each of these example check whether the given linear transformation is (i) onto, (ii) one-to-one and (iii) isomorphism.

Solution : In case of example (3), since the matrix $A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is non-singular, hence for $\vec{u} = (x_1, y_1)$ and $\vec{v} = (x_2, y_2)$ in $\mathbb{R}^2$, if $T_A(\vec{u}) = T_A(\vec{v})$, then

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

implies that $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$

That is $\vec{u} = \vec{v}$, so T_A is one-to-one.

Similarly one can see that T_A is onto also, thus T_A is an isomorphism.

In case of example (4), clearly range of T is $\{\vec{0}\}$; that is $T(V) = \{\vec{0}\}$ and therefore it is not onto as well as it not one-to-one.

For example (5) I is one-to-one and onto also, and so it is an isomorphism.

For example (6), if $k = 0$, then this a example (4), so in this case it is neither onto nor one-to-one. But if $k \neq 0$, then for any $\vec{u}, \vec{v}$ in V if $T(\vec{u}) = T(\vec{v})$, that is, $k\vec{u} = k\vec{v}$ which implies $k(\vec{u} - \vec{v}) = 0$, hence $\vec{u} - \vec{v} = 0$ as $k \neq 0$.

$\therefore \quad \vec{u} = \vec{v}$, showing that T is one-to-one. Also it is clear that T is onto. Thus in this case if $k \neq 0$, then T is an isomorphism of V onto V itself.

5.3 Sum, Scalar Multiple and Composition of Linear Transformations

Let T_1, T_2 be linear transformations of a vector space V into the vector space W. If α is any scalar, then

(i) $T_1 + T_2 : V \to W$ is defined by

$$(T_1 + T_2)(\vec{v}) = T_1(\vec{v}) + T_2(\vec{v}), \text{ for any } \vec{v} \in V.$$

(ii) $\alpha T_1 : V \to W$ defined by $(\alpha T_1)(\vec{v}) = \alpha [T_1(\vec{v})]$ are both linear maps; called sum and scalar multiple of linear maps.

We verify that $T_1 + T_2$ and αT_1 are linear maps.

(i) Let $\vec{u}, \vec{v}$ be any vectors in V, then

$$\begin{aligned}
[T_1 + T_2](\vec{u} + \vec{v}) &= T_1(\vec{u} + \vec{v}) + T_2(\vec{u} + \vec{v}) && \text{by (i)} \\
&= [T_1(\vec{u}) + T_1(\vec{v})] + [T_2(\vec{u}) + T_2(\vec{v})] \\
& && \because T_1 \text{ and } T_2 \text{ are linear maps} \\
&= [T_1(\vec{u}) + T_2(\vec{u})] + [T_1(\vec{v}) + T_2(\vec{v})] \\
& && \text{using vector space axioms.} \\
&= (T_1 + T_2)(\vec{u}) + (T_1 + T_2)(\vec{v}) && \text{using (i)}
\end{aligned}$$

Also for any scalar k,

$$\begin{aligned}
(T_1 + T_2)(k\vec{u}) &= T_1(k\vec{u}) + T_2(k\vec{u}), && \text{using (i)} \\
&= kT_1(\vec{u}) + kT_2(\vec{u}) && \text{as } T_1 \text{ and } T_2 \text{ are linear maps} \\
&= k[T_1(\vec{u}) + T_2(\vec{u})] \\
&= k(T_1 + T_2)(\vec{u})
\end{aligned}$$

Thus, $T_1 + T_2$ is a linear map.

(ii) For $\vec{u}$, $\vec{v}$ in V, we have,

$$(\alpha T_1)(\vec{u} + \vec{v}) = \alpha [T_1(\vec{u} + \vec{v})] \quad \text{using (ii)}$$

$$= \alpha [T_1(\vec{u}) + T_2(\vec{v})] \quad \because T_1 \text{ is linear map}$$

$$= \alpha T_1(\vec{u}) + \alpha T_1(\vec{v}) \quad \text{using vector space axioms}$$

$$= (\alpha T_1)(\vec{u}) + (\alpha T_1)(\vec{v}) \quad \text{using (ii)}$$

Thus αT_1 preserves addition.

Also, for any scalar k,

$$(\alpha T_1)(k\vec{u}) = \alpha [T_1(k\vec{u})] \quad \text{by (ii)}$$

$$= \alpha [k\, T_1(\vec{u})] \quad \text{as } T_1 \text{ is linear map}$$

$$= (\alpha k)\, T_1(\vec{u}) \quad \text{By axiom } M_3$$

$$= (k\alpha)\, T_1(\vec{u}) \quad \because \alpha k = k\alpha$$

$$= k [\alpha\, T_1(\vec{u})] \quad \text{by axiom } M_3$$

$$= k (\alpha T_1)(\vec{u}) \quad \text{by (ii) again}$$

Hence, αT_1 is a linear map.

Definition : Let V and W be vector spaces over $\mathbb{F}$. Then, the set of all linear transformations (maps) from V to W is denoted by L(V, W) or Hom (V, W).

Theorem 1 : If V and W are vector spaces over $\mathbb{F}$, then Hom(V, W) is also a vector space over $\mathbb{F}$, with operations of addition and scalar multiplication defined as above.

Proof : The closure axioms of addition and scalar multiplication follow from above (i) and (ii). Rest of the proof is routine and is left as an exercise for the students. One may refer chapter 1, where we have illustrated all steps in the example of **function space**.

Remark : If in particular, W is replaced by $\mathbb{R}$, then L (V, $\mathbb{R}$) is also a vector space and it is usually called the **dual space** of V. The elements of L (V, $\mathbb{R}$) are called **linear functions** or **linear forms**, thus any linear map $f : V \to \mathbb{R}$ is called linear form.

Definition : Let U, V and W be real vector spaces and $T_1 : U \to V$, $T_2 : V \to W$ be linear transformations. The *composition of T_2 with T_1*, denoted by $T_2 \circ T_1$ is a map defined by the formula :

$$(T_2 \circ T_1)(\vec{u}) = T_2 [T_1(\vec{u})] \quad \ldots (*)$$

where, $\vec{u}$ is a vector in U.

Theorem 2 : If $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations, then $T_2 \circ T_1 : U \to W$ is also a linear transformation.

Proof : If $\vec{u}$ and $\vec{v}$ are two vectors in U and k is any scalar, we have

$$(T_2 \circ T_1)(\vec{u} + \vec{v}) = T_2[T_1(\vec{u} + \vec{v})] \quad \text{by formula (*)}$$

$$= T_2[T_1(\vec{u}) + T_1(\vec{v})] \quad \text{by linearity of } T_1$$

$$= T_2(T_1(\vec{u})) + T_2(T_1(\vec{v})) \quad \text{by linearity of } T_2$$

$$= (T_2 \circ T_1)(\vec{u}) + (T_2 \circ T_1)(\vec{v}), \quad \text{by formula (*)}$$

Also $$(T_2 \circ T_1)(k\vec{u}) = T_2[T_1(k\vec{u})], \quad \text{by formula (*)}$$

$$= T_2[k\, T_1(\vec{u})] \quad \text{by homogeneity of } T_1$$

$$= k\, T_2[(T_1(\vec{u})] \quad \text{by homogeneity of } T_2$$

$$= k\,(T_2 \circ T_1)(\vec{u})$$

Thus, $T_2 \circ T_1$ satisfies the two requirements of a linear transformation.

SOLVED EXAMPLES

Example 5.14 : Let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^3$ be transformations given by,

$$T_1(x, y) = (x + y, y)$$

and $$T_2(x_1, y_1) = (2x_1, y_1, x_1 + y_1)$$

show that T_1 and T_2 are linear transformations and find formula for $T_2 \circ T_1$.

Solution : For $\vec{u} = (x, y)$ and $\vec{v} = (x', y')$ be any two vectors in $\mathbb{R}^2$, we have $(x, y) + (x' y') = (x + x', y + y')$

then $$T_1(\vec{u} + \vec{v}) = T(x + x', y + y')$$

$$= \big((x + x') + (y + y'), y + y'\big) \quad \text{by formula of } T_1$$

$$= \big((x + y) + (x' + y'), y + y'\big)$$

$$= (x + y, y) + (x' + y', y')$$

$$= T_1(x, y) + T_1(x', y')$$

$$= T_1(\vec{u}) + T_1(\vec{v})$$

Also, for any scalar k, $k\vec{u} = k(x, y) = (kx, ky)$

$$\begin{aligned} T_1(k\vec{u}) &= T_1(kx, ky) \\ &= (kx + ky, ky) \\ &= (k(x + y), ky) \\ &= k(x + y, y) \\ &= k\,T_1(x, y) \\ &= kT_1(\vec{u}) \end{aligned}$$

Similarly, one can check T_2 is a linear transformation, it is left as an exercise. Now, for $\vec{u} = (x, y)$ in R^2, we have

$$\begin{aligned} (T_2 \circ T_1)(\vec{u}) &= T_2[T_1(\vec{u})] \\ &= T_2(x + y, y) && \text{by formula of } T_1 \\ &= (2(x + y), y, x + y + y) \\ &= (2x + 2y, y, x + 2y) && [\text{Take } x_1 = x + y \text{ and } y_1 = y \text{ in formula of } T_2] \end{aligned}$$

which is formula for $T_2 \circ T_1$.

Example 5.15 : Let A be a 2×3 matrix and $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the matrix transformation $T(x) = AX$. given by :

$$T\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}, \quad T\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}3\\0\end{bmatrix}, \quad T\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}4\\-7\end{bmatrix}$$

Find (a) $T\begin{bmatrix}1\\-2\\3\end{bmatrix}$, (b) $T\begin{bmatrix}a\\b\\c\end{bmatrix}$, (c) A

Solution : (a) We have

$$\begin{bmatrix}1\\-2\\3\end{bmatrix} = (1)\begin{bmatrix}1\\0\\0\end{bmatrix} + (-2)\begin{bmatrix}0\\1\\0\end{bmatrix} + (3)\begin{bmatrix}0\\0\\1\end{bmatrix}$$

By linearity of T, we obtain

$$\begin{aligned} T\begin{bmatrix}1\\-2\\3\end{bmatrix} &= 1 \cdot T\begin{bmatrix}1\\0\\0\end{bmatrix} + (-2)\begin{bmatrix}0\\1\\0\end{bmatrix} + 3\,T\begin{bmatrix}0\\0\\1\end{bmatrix} \\ &= 1 \cdot \begin{bmatrix}1\\1\end{bmatrix} + (-2)\begin{bmatrix}3\\0\end{bmatrix} + 3\begin{bmatrix}4\\-7\end{bmatrix} \\ &= \begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}-6\\0\end{bmatrix} + \begin{bmatrix}12\\-21\end{bmatrix} = \begin{bmatrix}7\\-20\end{bmatrix} \end{aligned}$$

(b) Also, $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ we have

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = a\,T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) + b\,T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) + c\,T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$$

$$= a\begin{bmatrix} 1 \\ 1 \end{bmatrix} + b\begin{bmatrix} 3 \\ 0 \end{bmatrix} + c\begin{bmatrix} 4 \\ -7 \end{bmatrix}$$

$$= \begin{bmatrix} a \\ a \end{bmatrix} + \begin{bmatrix} 3b \\ 0 \end{bmatrix} + \begin{bmatrix} 4c \\ -7c \end{bmatrix}$$

$$= \begin{bmatrix} a + 3b + 4c \\ a + 0 + (-7)\,c \end{bmatrix} \quad \ldots (*)$$

(c) From (*)

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} a + 3b + 4c \\ a - 7c \end{bmatrix} = \begin{bmatrix} 1 & 3 & 4 \\ 1 & 0 & -7 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Thus, $T(X) = AX$, gives

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 1 & 0 & -7 \end{bmatrix}$$

5.4 Kernel and Range

In this section we shall discuss some properties of linear transformations and explore various relationships between linear transformations and matrices.

Theorem 3 : If $T : V \to W$ is a linear transformation, then

(a) $T(\vec{0}) = \vec{0}$ (note that $\vec{0}$ on R.H.S. is $\vec{0}$ in W)

(b) $T(-\vec{u}) = -T(\vec{u})$ for all $\vec{u}$ in V

(c) $T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v})$, for all $\vec{u}$ and $\vec{v}$ in V.

Proof : Let $\vec{u}$ be any vector in V. Since $0\vec{u} = \vec{0}$.

We have $T(\vec{0}) = T(0.\vec{u}) = 0.T(\vec{u}) = \vec{0}$
which proves (a).

Also $T(-\vec{u}) = T[(-1)\vec{u}] = (-1)\,T(\vec{u}) = -T(\vec{u})$
which proves (b).

Finally, $T(\vec{u} - \vec{v}) = T(\vec{u} + (-1)\vec{v}) = T(\vec{u}) + T((-1)\vec{v})$

$= T(\vec{u}) + (-1)\,T(\vec{v}) = T(\vec{u}) - T(\vec{v})$

which proves (c).

Definition : Let $T : V \to W$ be a linear transformation, then the set of all vectors in V that T maps into $\vec{0}$ is called the *Kernel (or null space)* of T, it is denoted by Ker (T).

Thus, $\text{Ker}(T) = \{\vec{u} \in V \mid T(\vec{u}) = \vec{0}\}$

The set of all vectors in W that are images under T of at least one vector in V is called the *range* of T, it is denoted by R(T).

Thus, $R(T) = \{\vec{w} \in W \mid \vec{w} = T(\vec{u}), \text{ for some } \vec{u} \text{ in } V\}$

Example 1 : Let V and W be real vector spaces.

Let $T : V \to W$ be zero linear transformation ; that is, $T(\vec{u}) = \vec{0}$ for all $\vec{u}$ in V. Then, obviously all vectors in V map into $\vec{0}$,

$\text{Ker}(T) = V$

Since only zero element in W is mapped by all elements in V under T, $R(T) = \{\vec{0}\}$.

Example 2 : Let V be any vector space and $I : V \to V$ be the identity transformation. Then, $\text{Ker}(I) = \{\vec{u} \in V \mid T(\vec{u}) = \vec{0}\}$ since, under identity map only zero vector in V will map into $\vec{0}$ of W,

$\text{Ker}(I) = \{\vec{0}\}$

Also, R(T) = V, since each element of V map to itself under T.

Theorem 4 : Let $T : V \to W$ be a linear transformation, then

(a) the kernel of T is a subspace of V

(b) the range of T is a subspace of W

Proof : (a) From (a) of theorem (2), since $T(\vec{0}) = \vec{0}$, $\vec{0}$ is in ker (T). This shows that ker (T) is not empty. Now, to show that ker (T) is subspace of V, we have to show that ker (T) is closed under addition and scalar multiplication. For $\vec{u}$ and $\vec{v}$ in ker (T), $T(\vec{u}) = \vec{0}$ and $T(\vec{v}) = \vec{0}$. Then $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) = \vec{0} + \vec{0} = \vec{0}$, showing that $\vec{u} + \vec{v}$ is in ker (T).

Also, for any scalar k,

$T(k\vec{u}) = kT(\vec{u}) = k.\vec{0} = \vec{0}$, showing that $k\vec{u}$ is also in ker (T).

Thus, ker (T) is a subspace of V.

(b) Let $\vec{w_1}$ and $\vec{w_2}$ be any vectors in R(T), then by definition there are $\vec{u_1}$ and $\vec{u_2}$ in V such that $T(\vec{u_1}) = \vec{w_1}$ and $T(\vec{u_2}) = \vec{w_2}$.

Since $\vec{u_1}$ and $\vec{u_2}$ are in V, $\vec{u_1} + \vec{u_2}$ is in V so that $T(\vec{u_1} + \vec{u_2})$ is in R(T).

But $T(\vec{u_1} + \vec{u_2}) = T(\vec{u_1}) + T(\vec{u_2})$, by linearity of T

$= \vec{w_1} + \vec{w_2}$

Thus, $T(\vec{u_1} + \vec{u_2}) = \vec{w_1} + \vec{w_2}$ is in R(T).

Also, for any scalar k, $k\vec{u_1}$ is in V, hence $T(k\vec{u_1})$ in R(T).

But, $T(k\vec{u_1}) = kT(\vec{u_1})$

$= k\vec{w_1}$

showing that $k\vec{w_1}$ is in R(T).

Thus, R(T) is a subspace of W.

SOLVED EXAMPLES

Example 5.16 : Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by

$T(x, y) = (x, x + y, y)$.

Find the range and kernel of T.

Solution : $\ker(T) = \{(x, y) \in \mathbb{R}^2 \mid T(x, y) = (0, 0, 0)\}$

Now, $T(x, y) = (0, 0, 0)$ if and only if $(x, x + y, y) = (0, 0, 0)$

i.e. $x = 0, \quad x + y = 0, \quad y = 0.$

Solving this we get $x = 0$ and $y = 0$.

This means only $(0, 0) \in \mathbb{R}^2$ maps to $(0, 0, 0)$ in $\mathbb{R}^3$. Thus $\ker T = \{(0, 0)\}$. Hence, $\dim \ker T = 0$.

Suppose $(x, y, z) \in$ Im (T), then there exists $(x_1, y_1) \in \mathbb{R}^2$ such that $T(x_1, y_1) = (x, y, z)$ or

$$(x_1, x_1 + y_1, y_1) = (x, y, z)$$

This gives $x = x_1, \; x_1 + y_1 = y, \; y_1 = z$.

Solving for x, y, z in terms of x_1, y_1, we obtain $x = x_1, y = x_1 + y_1$ and $z = y_1$.

This shows that $y = x + z$.

$\therefore$ $(x, y, z) \in$ Im (T) only if $y = x + z$, hence

$$\text{Im}(T) = \{(x, x + z, z) \mid x, z \in \mathbb{R}\}$$

$$= \{x(1, 1, 0) + z(0, 1, 1) \mid x, z \in \mathbb{R}\}$$

This implies that $\dim \text{Im}(T) = 2$ and basis of Im (T) is $\{(1, 1, 0), (0, 1, 1)\}$.

Example 5.17 : $T : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$T(x, y, z) = (x + y + 2z, x + z, 2x + y + 3z)$$

Find kernel and image of T.

Solution : $\ker T = \{(x, y, z) \in \mathbb{R}^3 \mid T(x, y, z) = (0, 0, 0)\}$

Now, $T(x, y, z) = (0, 0, 0)$ if and only if $(x + y + 2z, x + z, 2x + y + 3z) = (0, 0, 0)$

i.e.
$$\begin{aligned} x + y + 2z &= 0 \\ x + z &= 0 \\ 2x + y + 3z &= 0 \end{aligned}$$

Second equation gives $z = -x$ and then using this in first equation or last equation we obtain $y = x$.

Therefore, $\ker(T) = \{(x, x, -x) \mid x \in \mathbb{R}\}$

$= \{x(1, 1, -1) \mid x \in \mathbb{R}\}$

dim ker (T) = 1 and basis for ker (T) is $\{(1, 1, -1)\}$.

Next let $(x, y, z) \in \mathbb{R}^3$. If $(x, y, z) \in$ Im (T), then there exists $(x_1, y_1, z_1) \in \mathbb{R}^3$ such that

$$\begin{aligned} T(x_1, y_1, z_1) &= (x, y, z) \text{ or} \\ x_1 + y_1 + 2z_1 &= x \\ x_1 + z_1 &= y \\ 2x_1 + y_1 + 3z_1 &= z \end{aligned}$$

This shows that $x - 2y = y_1 - x_1$ and $z - 3y = y_1 - x_1$. Therefore, $x - 2y = z - 3y$, this gives $x + y = z$. Thus, $(x, y, z) \in$ Im (T) only if $x + y = z$. Hence,

$$\begin{aligned} \text{Im}(T) &= \{(x, y, x + y) \mid x, y \in \mathbb{R}\} \\ &= \{x(1, 0, 1) + y(0, 1, 1) \mid x, y \in \mathbb{R}\} \end{aligned}$$

$\therefore$ dim Im (T) = 2 and basis of Im (T) is $\{(1, 0, 1), (0, 1, 1)\}$.

Theorem 5 : If $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by an $m \times n$ matrix A, then

(a) The kernel of T is the null space of A ;

(b) The range of T is the column space of A.

Proof : Left as an exercise for the students use (Example (3) and (4) for general presentation of the proof).

Definition : If $T : V \to W$ is a linear transformation then the dimension of the range of T is called the *rank of T* and is denoted by rank (T), the dimension of the kernel of T is called the *nullity of T* and is denoted by nullity (T).

Note : If $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by an $m \times n$ matrix A, then we have two definitions of rank and nullity, one for A and one for T. The following theorem shows that those definitions do not conflict.

Theorem 6 : If A is an $m \times n$ matrix and $T : \mathbb{R}^n \to \mathbb{R}^m$ is multiplication by A, then

(a) nullity (T) = nullity (A)

(b) rank (T) = rank (A)

Proof : From theorem 4 and definitions of rank and nullity for T and A we have

$$\begin{aligned} \text{nullity (A)} &= \text{dim (null space of A)} \\ &= \text{dim [ker (T)]} \\ &= \text{nullity (T).} \\ \text{rank (A)} &= \text{dim (column space of A)} \\ &= \text{dim [R(T)]} \\ &= \text{rank (T).} \end{aligned}$$

Now, we state *Dimension Theorem for linear transformations.*

Theorem 7 : (Dimension Theorem for linear Transformations)

If $T : V \to W$ is a linear transformation from an n-dimensional vector space V to a vector space W, then

$$\text{rank (T)} + \text{nullity (T)} = n.$$

In other words, for linear transformation the rank plus the nullity is equal to the dimension of the domain.

Proof : Let $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}\}$ be a basis of ker (T) where, r = dim [ker (T)] ≤ n; since ker (T) is subspace of V; that is, nullity (T) = r ≤ n. Since $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}\}$ is basis for a subspace of V, it is linearly independent, so can be extended to a basis

$$B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}, \vec{u_{r+1}} \ldots, \vec{u_{r+k}}\}$$

for V, where $r + k = n$ ($\because$ dim V = n)

Note that $r = \text{nullity (T)}$.

We have to prove that the k elements in the set

$$S = \{T(\vec{u_{r+1}}), T(\vec{u_{r+2}}), \ldots, T(\vec{u_{r+k}})\}$$

form a basis for R(T) ; that is, dim R(T) = k or rank (T) = r.

First let us show that S spans R(T). For this, let $\vec{w}$ be any element in R(T), then by definition of R(T), there is $\vec{u}$ in V such that $T(\vec{u}) = \vec{w}$.

Since B is a basis for V, $\vec{u}$ can be expressed as a linear combination of elements B, as

$$\vec{u} = \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_r \vec{u_r} + \beta_1 \vec{u_{r+1}} + \beta_2 \vec{u_{r+2}} + \ldots + \beta_k \vec{u_{r+k}} \quad \ldots (1)$$

where $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_k$ are scalars.

Applying T on both sides of (1) we have

$$\vec{w} = T(\vec{u}) = T(\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_r \vec{u_r} + \beta_1 \vec{u}_{r+1} + \ldots + \beta_k \vec{u}_{r+k})$$

$$= \alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \ldots + \alpha_r T(\vec{u}_r) + \beta_1 T(\vec{u}_{r+1})$$

$$+ \beta_2 T(\vec{u}_{r+2}) + \ldots + \beta_k T(\vec{u}_{r+k}), \qquad \text{by linearity of T.}$$

$$= \beta_1 T(\vec{u}_{r+1}) + \beta_2 T(\vec{u}_{r+2}) + \ldots + \beta_k T(\vec{u}_{r+k})$$

since $T(\vec{u_1}) = T(\vec{u_2}) = \ldots\ldots = T(\vec{u_r}) = 0$ as $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}$ are in ker (T).

Thus, $\vec{w}$ is linear combination of elements in S, hence S spans R(T).

Now, to show that S is linearly independent, suppose we have, scalars $\beta_1, \beta_2, \ldots, \beta_k$ such that

$$\beta_1 T(\vec{u}_{r+1}) + \beta_2 T(\vec{u}_{r+2}) + \ldots + \beta_K T(\vec{u}_{r+k}) = 0 \qquad \ldots (2)$$

By linearity of T, this becomes

$$T(\beta_1 \vec{u}_{r+1} + \beta_2 \vec{u}_{r+2} + \ldots + \beta_k \vec{u}_{r+k}) = 0$$

The last equation shows that the element

$$\vec{u} = \beta_1 \vec{u}_{r+1} + \beta_2 \vec{u}_{r+2} + \ldots + \beta_k \vec{u}_{r+k} \qquad \ldots (3)$$

is in ker (T). But ker (T) is spanned by $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}\}$. Hence $\vec{u}$ must also be a linear combination of $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_r}$, say

$$\vec{u} = \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_r \vec{u_r}$$

From (2) and (3) we obtain

$$\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_r \vec{u_r} + (-\beta_1) \vec{u}_{r+1} + \ldots + (-\beta_k) \vec{u}_{r+k} = 0 \qquad \ldots (5)$$

But $B = \{\vec{u_1}, \vec{u_2}, \ldots \vec{u_r}, \vec{u}_{r+1}, \ldots, \vec{u}_{r+k}\}$ is basis for V, hence is linearly independent, so that equation (5) yields,

$$\alpha_1 = \alpha_2 = \ldots\ldots = \alpha_r = 0 = \beta_1 = \beta_2 = \ldots\ldots = \beta_k,$$

that is $\beta_1 = \beta_2 = \ldots\ldots = \beta_k = 0$

This factor with (2) shows that S is linearly independent. Thus, S spans R(T) and is linearly independent. Hence S is a basis for R(T), that is,

$$\dim R(T) = \text{rank } (T) = k$$

Thus, we have

$$n = r + k = \text{nullity } (T) + \text{rank } (T)$$

Theorem 8 : (Linear Transformation with prescribed values)

If $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ is a basis for V and $\vec{w_1}, \vec{w_2}, \ldots, \vec{w_n}$ are arbitrary n vectors in W, not necessarily distinct, then there exists a linear transformation $T : V \to W$ such that $T(\vec{u_i}) = \vec{w_i}$, i = 1, 2, … n.

Proof : Let $\vec{u}$ be any vector in V. Since $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ is basis for V, there exists unique set of scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\vec{u} = \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_n \vec{u_n} \quad \ldots (1)$$

In fact, here $\alpha_1, \alpha_2 \ldots, \alpha_n$ are co–ordinates of $\vec{u}$ relative to the basis B. This suggests us to define a map $T : V \to W$ as

$$T(\vec{u}) = \alpha_1 \vec{w_1} + \alpha_2 \vec{w_2} + \ldots + \alpha_n \vec{w_n} \quad \ldots (2)$$

We shall prove that T is a linear transformation by establishing that T preserves addition and scalar multiplication. Let $\vec{u}$ and $\vec{v}$ be any vectors in V and suppose $\vec{u}$ has representation relative to B as in (1) and

$$\vec{v} = \beta_1 \vec{u_1} + \beta_2 \vec{u_2} + \ldots + \beta_n \vec{u_n}$$

Then $\vec{u} + \vec{v} = (\alpha_1 + \beta_1) \vec{u_1} + (\alpha_2 + \beta_2) \vec{u_2} + \ldots + (\alpha_n + \beta_n) \vec{u_n}$

By definition of T, we have

$$\begin{aligned} T(\vec{u} + \vec{v}) &= (\alpha_1 + \beta_1) \vec{w_1} + (\alpha_2 + \beta_2) \vec{w_2} + \ldots + (\alpha_n + \beta_n) \vec{w_n} \\ &= (\alpha_1 \vec{w_1} + \alpha_2 \vec{w_2} + \ldots + \alpha_n \vec{w_n}) + (\beta_1 \vec{w_1} + \beta_2 \vec{w_2} + \ldots + \beta_n \vec{w_n}) \\ &= T(\vec{u}) + T(\vec{v}) \end{aligned}$$

Also, for any scalar k

$$k\vec{u} = (k\alpha_1) \vec{u_1} + (k\alpha_2) \vec{u_2} + \ldots + (k\alpha_n) \vec{u_n}$$

Therefore again by definition of T,

$$\begin{aligned} T(k\vec{u}) &= (k\alpha_1) \vec{w_1} + (k\alpha_2) \vec{w_2} + \ldots + (k\alpha_n) \vec{w_n} \\ &= k(\alpha_1 \vec{w_1} + \alpha_2 \vec{w_2} + \ldots + \alpha_n \vec{w_n}) \\ &= kT(\vec{u}) \end{aligned}$$

Thus T is a linear transformation. Now, for each i, since

$$\vec{u_i} = 0.\vec{u_1} + 0.\vec{u_2} + \ldots + 1\vec{u_i} + \ldots + 0.\vec{u_n}$$

$$T(\vec{u_i}) = 0.\vec{w_1} + 0.\vec{w_2} + \ldots + 1\vec{w_i} + \ldots + 0.\vec{w_n}$$

$$= \vec{w_i}$$

Showing that $T(\vec{u_i}) = \vec{w_i}$; i = 1, 2, … , n. Thus, T is a linear transformation satisfying the given condition.

Note : The linear transformation obtained in above theorem is unique, since the co-ordinates of any vector in V are uniquely determined. In fact, suppose there are two linear transformations T and T' satisfying the condition given in the theorem, that is $T(\vec{u_i}) = \vec{w_i} = T'(\vec{u_i})$, i = 1, 2, … n. Then for $\vec{u}$ in V,

$$T(\vec{u}) = \alpha_1 \vec{w_1} + \alpha_2 \vec{w_2} + \ldots + \alpha_n \vec{w_n}$$

$$= \alpha_1 T'(\vec{u_1}) + \alpha_2 T'(\vec{u_2}) + \ldots + \alpha_n T'(\vec{u_n})$$

$$= T'(\alpha_1 \vec{u_1}) + T'(\alpha_2 \vec{u_2}) + \ldots + T'(\alpha_n \vec{u_n})$$

$$= T'(\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_n \vec{u_n}) = T'(\vec{u})$$

Thus, $T(u) = T'(\vec{u})$, for all $\vec{u}$ in V.

$\therefore \quad T = T'$

Remark : The above theorem also states that a linear transformation is completely determined by its action on a basis of a vector space V.

Let us prove one more result.

Theorem 9 : Let T : V → W be a linear transformation and $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be a basis for V. Then show that range of T, R(T) is spanned by the set $B' = \{T(\vec{u_1}), T(\vec{u_2}), \ldots, T(\vec{u_n})\}$

Proof : Let $\vec{w}$ be any vector in R(T), then by definition of R(T), there is $\vec{u}$ in V such that $T(\vec{u}) = \vec{w}$.

Since B is basis for V, B spans V, hence for $\vec{u}$ in V, we have

$$\vec{u} = \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_n \vec{u_n} .$$

Applying T on both sides and using linearity of T we get

$$T(\vec{u}) = \alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \ldots + \alpha_n T(\vec{u_n})$$

which shows that $T(\vec{u})$ is a linear combination of vectors in B'. Thus B' spans R(T).

Important note : The vectors $T(\vec{u_1})$, $T(\vec{u_2})$, … , $T(\vec{u_n})$ **need not be linearly independent**, need not be distinct.

For example, if $T : \mathbb{R}^3 \to \mathbb{R}^3$ is linear transformation determined by

$$T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\end{bmatrix}$$

$$T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}-2\\-4\\-6\end{bmatrix} \text{ and } T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

Since the vectors $\vec{u_1} = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\vec{u_2} = \begin{bmatrix}0\\1\\0\end{bmatrix}$ and $\vec{u_3} = \begin{bmatrix}0\\0\\1\end{bmatrix}$ is a basis for $\mathbb{R}^3$, by theorem 7, T is a linear transformation and is uniquely determined. But the vectors $T(\vec{u_1})$, $T(\vec{u_2})$ and $T(\vec{u_3})$; that is $\begin{bmatrix}1\\2\\3\end{bmatrix}$, $\begin{bmatrix}-2\\-4\\-6\end{bmatrix}$ and $\begin{bmatrix}0\\0\\0\end{bmatrix}$ are not linearly independent, which can be verified easily.

Theorem 10 : Let V be a vector space of dimension n and W be a vector space of dimension m, then the vector space L(V, W) the set of all linear transformations from V into W, has dimension mn, that is

$$\dim L(V, W) = \dim V \times \dim W$$

Proof : Let $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be a basis for V and $\{\vec{w_1}, \vec{w_2}, \ldots, \vec{w_m}\}$ be that of W. In view of the theorem 8, there exists unique linear transformation $T_{ij} : V \to W$ in L(V, W) defined by

$$T_{ij}(\vec{u_k}) = \vec{0} \quad \text{for } k \neq i$$

$$T_{ij}(\vec{u_i}) = \vec{w_j}, \quad i = 1, 2, \ldots, n, j = 1, 2, \ldots, m \qquad \ldots (i)$$

Thus, we see that there are uniquely defined mn linear transformations $\{T_{ij}\}_{i = 1 \text{ to } n,\ j = 1 \text{ to } m}$ from V into W. We claim that these mn linear transformations form a basis of L(V, W), for this we have to prove that (a) $T_{ij's}$ span the vector space L(V, W) and (b) $T_{ij's}$ are linearly independent.

Let $T \in L(V, W)$ be any linear transformation from V to W. Then for each i = 1 to n the elements $T(\vec{u_i})$ in W can be expressed as

$$T(\vec{u_i}) = \sum_{j=1}^{m} \alpha_{ij} \vec{v_j} \qquad \ldots (ii)$$

We define a linear transformation in L(V, W) by

$$T_1 = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} T_{ij} \qquad \ldots(iii)$$

Now using equation (i) and (ii), we have

$$T_1(\vec{u_k}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} T_{ij} \vec{u_k}$$

$$= \sum_{j=1}^{m} \alpha_{kj} T_{kj} \vec{u_k} \qquad (\because T_{ij}(\vec{u_k}) = \vec{0} \text{ for } i \neq k \text{ for (i)})$$

$$= \sum_{j=1}^{m} \alpha_{kj} \vec{v_i} \qquad (\because T_{ki}(\vec{u_k}) = \vec{v_j}, \text{ by (i)})$$

$$= T(\vec{u_k}) \qquad \ldots (iv)$$

Thus T and T_1 agree on each basis elements, hence $T(\vec{u}) = T_1(\vec{u})$, for all $\vec{u} \in V$, so $T = T_1$. This shows that any linear transformation T in L(V, W) can be expressed as a linear combination of $T_{ij's}$.

To prove (b), by virtue of (iii), $T_1 = 0$ implies that

$$\sum_{j=1}^{m} \alpha_{kj} \vec{v_j} = \vec{0}$$

From this equation it follows that $\alpha_{kj} = 0$, k = 1, 2, …, n and j = 1, 2, …, m, since $\vec{v_{j's}}$ are linearly a independent. Therefore, thus $\{T_{ij}\}_{i=1 \text{ to } n,\ j=1 \text{ to } m}$ form a basis of L(V, W).

SOLVED EXAMPLES

Example 5.18 : Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation given by the formula $T(x, y, z) = (x + y - z, x - 2y + z, -2x - 2y + 2z)$.

(a) Which of the following given vectors are in ker (T) ?
(i) (1, 2, 3), (ii) (1, 2, 1), (iii) (– 1, 1, 2)

(b) Which of the following given vectors are in R(T) ?
(i) (1, 2, – 2), (ii) (3, 5, 2), (iii) (– 2, 3, 4)

Solution : (a) If X = (x, y, z) is in ker (T), then by definition T(X) = 0;
that is, $(x + y - z, x - 2y + z, -2x - 2y + 2z) = (0, 0, 0)$
or, equating corresponding components,

$$\begin{aligned} x + y - z &= 0 \\ x - 2y + z &= 0 \\ -2x - 2y + 2z &= 0 \end{aligned} \qquad \ldots (*)$$

Therefore, if X = (x, y, z) is in ker (T), its co-ordinates x, y and z must satisfy the system (*).

It is easy to check that (i) (1, 2, 3) and (iii) (– 1, 1, 2) satisfy (*) and hence are in ker (T), whereas (ii) (1, 2, 1) cannot satisfy last equation in (*), hence cannot be in ker (T).

(b) In order to have $\vec{u}$ = (a, b, c) to be in R(T), we must find $\vec{x}$ = (x, y, z) in $\mathbb{R}^3$ such that

$$T(\vec{x}) = (a, b, c)$$

or $(x + y - z, x - 2y + z, -2x - 2y + 2z) = (a, b, c)$

or equating corresponding components,

$$\begin{aligned} x + y - z &= a \\ x - 2y + z &= b \\ -2x - 2y + 2z &= c \end{aligned} \quad \ldots (**)$$

Thus, if (a, b, c) is in R(T), there should be $\vec{x}$ = (x, y, z) in R^3, which satisfies the system of linear equations in (* *).

The system (* *) can be rewritten, by adding two times first equation in third equation as

$$\begin{aligned} x + y - z &= a \\ x - 2y + z &= b \\ 0 &= 2a + c \end{aligned}$$

The last equation, that is, the identity 2a + c = 0, implies that the system in (**) is consistent if and only if 2a + c = 0.

In given vectors (i) (1, 2, – 2) and (ii) (– 2, 3, 4) satisfy this condition, whereas the vector (iii) (3, 5, 2) does not satisfy 2a + c = 0. Hence it cannot be in R(T).

Thus, it is sure that (1, 2, 1) cannot be in R(T). For confirmation of vectors (i) and (iii) to be in R(T), we proceed to solve the system (**) further, and we can obtain

$$x = \frac{2}{3}a + \frac{b}{3} + \frac{z}{3},$$

$$y = \frac{a}{3} - \frac{b}{3} + \frac{2}{3}z.$$

For given values of a and b; x, y, z exist satisfying (**), hence are in R(T).

Example 5.19 : If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation given by

$$T(x, y, z) = (x + y - z, x - 2y + z, -2x - 2y + 2z)$$

Find basis and dimension of (a) ker (T), (b) R(T).

Solution : (a) If $\vec{x}$ = (x, y, z) is in ker (T), its co-ordinates must satisfy the system of linear equations :

$$\begin{aligned} x + y - z &= 0 \\ x - 2y + z &= 0 \\ -2x - 2y + 2z &= 0 \end{aligned} \quad \ldots (*)$$

or, by adding two times first equation in third,

$$\begin{aligned} x + y - z &= 0 \\ x - 2y + z &= 0 \\ 0 &= 0 \end{aligned}$$

or, subtracting first equation from second

$$\begin{aligned} x + y - z &= 0 \\ -3y + 2z &= 0 \end{aligned}$$

The solution of this system is

$$x = \frac{t}{3} \; ; \; y = \frac{2}{3} t$$

$$z = t, \; t \in \mathbb{R}$$

$$\therefore \quad \ker (T) = \left\{ \left(\frac{1}{3} t, \frac{2}{3} t, t \right) / \; t \in \mathbb{R} \right\}$$

$$= \left\{ t \left(\frac{1}{3}, \frac{2}{3}, 1 \right) / \; t \in \mathbb{R} \right\}$$

$\therefore$ dim ker (T) = 1 and basis for it is $\left\{ \left(\frac{1}{3}, \frac{2}{3}, 1 \right) \right\}$.

Now, by rank-nullity theorem dim V = rank (T) + nullity (T)

Here, dim V = 3, and nullity (T) = 1

$\therefore$ rank (T) = 2.

But, to find the basis for R(T) we need to apply some method, which gives basis as well as dimension of R(T). For this, one of the ways is to consider T in matrix notations as

$$T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} x + y - z \\ x - 2y + z \\ -2x - 2y + 2z \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & 1 \\ -2 & -2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Thus, $T(X) = AX$, where,

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & 1 \\ -2 & -2 & 2 \end{bmatrix}$$

We know that, if B is in $\mathbb{R}^3$, there is $\vec{x}$ in $\mathbb{R}^3$ such that $T(\vec{x}) = A\vec{x} = B$, which shows that to each B in $\mathbb{R}^3$, AX = B has a solution ; that is, AX = B is consistent. We know that is the column space of A. So we need to find the column space of A; for this we consider A^t and reduce it to row–echelon form so that column space of A is the row space of A^t. We have

$$A^t = \begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & -2 \\ -1 & 1 & 2 \end{bmatrix}$$

The row–echelon form of A^t is

$$\begin{bmatrix} 1 & 1 & -2 \\ 0 & -3 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

or $$\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, the basis for row space of A^t is $\{(1, 1, -2), (0, 1, 0)\}$ and dimension of row–space is 2. Converting this to A, we have, the basis for column space of A is $\left\{\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$ and dimension of column space of A is 2. That is *rank (T) = 2.*

Example 5.20 : Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation such that $T(1, 0, 0) = (0, 0)$, $T(0, 1, 0) = (1, 1)$ and $T(0, 0, 1) = (1, -1)$. Compute $T(4, -1, 1)$ and determine the nullity and rank of T.

Solution : For any (a, b, c) in $\mathbb{R}^3$, we have

$$(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)$$

Since T is linear transformation

$$\begin{aligned} T(a, b, c) &= a\,T(1, 0, 0) + bT(0, 1, 0) + cT(0, 0, 1) \\ &= a(0, 0) + b(1, 1) + c(1, -1) \\ &= (b + c, b - c) \qquad \ldots (*) \end{aligned}$$

$$\therefore T(4, -1, 1) = (-1 + 1, -1 - 1) = (0, -2)$$

Now, replacing (a, b, c) by (x, y, z) and writing the formula (*) of T in matrix notation, we have

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y + z \\ y - z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \ldots (**)$$

Thus, $T(X) = AX$,

where, $A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

We know nullity(T) = nullity (A)

and rank (T) = rank (A)

One can work out easily that

rank (A) = 2

and nullity (A) = 1

Note : The null space of A is the solution space of the system

$$\begin{aligned} y + z &= 0 \\ y - z &= 0 \end{aligned}$$

which shows that y = 0 and z = 0, and is independent of x; that is x is free variable, hence dim (nulls pace of A) = 1 and basis is $\{(1, 0, 0)\}$.

Example 5.21 : Comment on the statement : If V and W are vector spaces and $T : V \to W$ is a linear transformation and $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}\}$ is a linearly independent set of vectors in V, then $\{T(\overrightarrow{v_1}), T(\overrightarrow{v_2}), \ldots, T(\overrightarrow{v_n})\}$ is also linearly independent in W.

Solution : The statement is not **true**, see the note followed by theorem 9, since $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is linearly independent in $\mathbb{R}^3$, where as their images (1, 2, 3), (–2, –4, –6), (0, 0, 0) under T are **not** linearly independent.

Example 5.22 : Can we construct a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^4$ such that

$$\text{Im } (T) = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_2 + x_3 + x_4 = 0\} ?$$

Solution : No. Because, observe that dim Im (T) = 3; that is rank T = 3. By rank-nullity theorem we have

$$2 = \dim \mathbb{R}^2 = \dim \ker T + \dim \text{Im } (T) > 3,$$

which is impossible.

Example 5.23 : Can we construct a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ such that

$$\text{Im } (T) = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\} ?$$

Solution : Yes. Here dim I_m (T) = 2 and rank-nullity theorem states that

$$2 = \dim \mathbb{R}^2 = \text{nullity } (T) + \text{rank } T$$
$$2 = \text{nullity } T + 2$$

Our T should be such that nullity (T) = 0 that is kernel of $T = \{\overrightarrow{0}\}$, that is T should be 1 – to – 1 linear transformation.

Thus at this state looking at the example 2 and theorem 11, we have proved.

Theorem 11 : Given any m × n matrix with real entries (i) $M_{n \times n}(\mathbb{R})$ there corresponds a linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ from $\mathbb{R}^n$ to $\mathbb{R}^m$. Conversely to each linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ there corresponds an m × n matrix [T] – the matrix representation of T with respect to some basis, not necessarily with respect to standard basis.

Note : (1) The corresponds from $M_{m \times n}(\mathbb{R})$ to hom $(\mathbb{R}^n, \mathbb{R}^m)$:

$$\psi : M_{m \times n}(\mathbb{R}) \text{ — } \text{Hom } (\mathbb{R}^n, \mathbb{R}^m)$$
$$: A \longrightarrow T_A$$

is 1-to-1, for the matrices with m = n = 2.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$$

$\psi(A) = \psi(B)$ if and only if

$AX = BX$, for all $x \in \mathbb{R}^2$

In particular if $X = \begin{pmatrix}1\\0\end{pmatrix}$, $Y = \begin{pmatrix}0\\1\end{pmatrix}$, we must have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and $$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This means $$\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} a' \\ c' \end{bmatrix} \text{ and } \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} b' \\ d' \end{bmatrix}$$

which gives $a' = a,\ b' = b,\ c' = c$ and $d' = d$.

$\therefore$ ψ is 1-to-1 correspondence.

(2) The discussion which is limited to $\mathbb{R}^n$ and $\mathbb{R}^m$ in above discussion will be extend to slightly general vector spaces V and W of dimensions n and m respectively over reals.

Exercise 5.1

1. Determine in each case of the following whether a map $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation ; if not give reason.
   (a) $T(x, y) = (2x, y)$
   (b) $T(x, y) = (x, y^2)$
   (c) $T(x, y) = (-y, x)$
   (d) $T(x, y) = (0, y)$
   (e) $T(x, y) = (x - y, 2x + y)$
   (f) $T(x, y) = (y, x + 1)$
   (g) $T(x, y) = (|x|, |y|)$
   (h) $T(x, y) = \left(\sqrt[3]{x}, \sqrt[3]{y}\right)$
2. In each case of the following formula given for a map $T : \mathbb{R}^3 \to \mathbb{R}^3$, determine whether T is a linear transformation.
   (a) $T(x, y, z) = (x, x + y, x + y, z)$
   (b) $T(x, y, z) = (1, 1, 1)$
   (c) $T(x, y, z) = (0, 0, 0)$
   (d) $T(x, y, z) = (x + 1, y + 1, z - 1)$
   (e) $T(x, y, z) = (x, 2y, 3z)$
   (f) $T(x, y, z) = (e^x, e^y, 0)$
3. Determine whether a map $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ given in each of the following is a linear transformation.
   (a) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = b + c$
   (b) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix}$
   (c) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = 3a - 4b + c - d$
   (d) $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a^2 + b^2$

4. Let B be 2×4 matrix and $T : M_{2 \times 2}(\mathbb{R}) \to M_{2 \times 4}(\mathbb{R})$ be a map given by $T(A) = AB$. Show that T is a linear transformation.
5. In each part, formulae are given for linear transformations $T_1 : \mathbb{R}^2 \to \mathbb{R}^3$ and $T_2 : \mathbb{R}^3 \to \mathbb{R}^2$. State the domain and image space of $T_2 \circ T_1$ and $T_1 \circ T_2$. Find formula for $(T_2 \circ T_1)(x, y)$.
   (a) $T_1(x, y) = (2x, -3y, x + y)$, $T_2(x, y, z) = (x - y, y + z)$
   (b) $T_1(x, y) = (x - y, y + x, x)$, $T_2(x, y, z)) = (0, x + y + z)$
6. Show that $T_1 \circ T_2 \neq T_2 \circ T_1$ for the linear operators $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ given by the formulae.
   $$T_1(x, y) = (y, x) \text{ and } T_2(x, y) = (0, x)$$
7. Let $\vec{u_1}$, $\vec{u_2}$ and $\vec{u_3}$ be vectors in a vector space V and $T : V \to \mathbb{R}^3$ a linear transformation for which $T(\vec{u_1}) = (2, -1, 2)$, $T(\vec{u_2}) = (3, 0, 2)$ and $T(\vec{u_3}) = (-2, 1, 3)$. Find $T(4\vec{u_1} + \vec{u_2} - 3\vec{u_3})$.
8. (a) Prove that if a, b, c and d are any scalars, then the formula $T(x, y) = (ax + by, cx + dy)$ defines linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$.
   (b) Does the formula $T(x, y) = \left((ax + by)^2, (cx + dy)^2\right)$ define a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$? Explain.
9. Prove that if $T : V \to W$ is a linear transformation,
   $$T(\vec{u} - \vec{v}) = T(\vec{u}) - T(\vec{v}) \text{ for all vectors } \vec{u} \text{ and } \vec{v} \text{ in V.}$$
10. Let $\left\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\right\}$ be a basis for a vector space V and $T : V \to W$ is a linear transformation.
    (a) Show that T is zero transformation if
    $$T(\vec{u_1}) = T(\vec{u_2}) = \ldots T(\vec{u_n}) = \vec{0}.$$
    (b) Show that T is identity transformation if
    $$T(\vec{u_1}) = \vec{u_1}, T(\vec{u_2}) = \vec{u_2}, \ldots, T(\vec{u_n}) = \vec{u_n}$$
11. If $\vec{e_1} = (1, 0, 0)$, $\vec{e_2} = (0, 1, 0)$ and $\vec{e_3} = (0, 0, 1)$ and if $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation such that
    $$T(\vec{e_3}) = 2\vec{e_1} + 3\vec{e_2} + 5\vec{e_3}, T(\vec{e_2} + \vec{e_3}) = \vec{e_1}$$
    $$T(\vec{e_1} + \vec{e_2} + \vec{e_3}) = \vec{e_2} - \vec{e_3}$$
    Compute $T(\vec{e_1} + 2\vec{e_2} + 3\vec{e_3})$ and determine the rank and nullity of T.

12. Let $\overrightarrow{e_1} = (1, 0)$ and $\overrightarrow{e_2} = (0, 1)$ be the basis vectors for $\mathbb{R}^2$ and $T : R^2 \to R^3$ map the basis vectors as follows : $T(\overrightarrow{e_1}) = (1, 0, 1)$ and $T(\overrightarrow{e_2}) = (-1, 0, 1)$. Compute $T(2\overrightarrow{e_1} - 2\overrightarrow{e_2})$ & determine the nullity and rank of T.
13. Solve the exercise (12) if $T(\overrightarrow{e_1}) = (1, 0, 1)$ and $T(\overrightarrow{e_2}) = (1, 1, 1)$.
14. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by
    $T(x, y) = (2x - y, -8x + 4y)$.
    (a) Which of the following vectors are in R(T) ?
        (i) (1, – 4), (ii) (5, 0), (iii) (– 3, 12).
    (b) Which of the following vectors are in ker (T) ?
        (i) (5, 10), (ii) (3, 2), (iii) (1, 1).
15. Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be the linear transformation given by the formula :
    $T(x_1, x_2, x_3, x_4) = (4x_1 + x_2 - 2x_3 - 3x_4, 2x_1 + x_2 + x_3 - 4x_4, 6x_1 - 6x_3 + 9x_4)$
    (a) Which of the following are in R(T) ?
        (i) (0, 0, 6), (ii) (1, 3, 0), (iii) (2, 4, 1).
    (b) Which of the following are in ker (T) ?
        (i) (3, – 8, 2, 0), (ii) (0, 0, 0, 1), (iii) (0, – 4, 1, 0).
16. Find a basis for the kernel of
    (a) linear transformation given in exercise (14)
    (b) the linear transformation in exercise (15).
17. Find a basis for the range of :
    (a) linear transformation given in exercise (14).
    (b) linear transformation given in exercise (15).
18. Verify rank nullity theorem for :
    (a) linear transformation given in exercise (14).
    (b) linear transformation given in exercise (15).

    Through exercises 19 – 23 let T be multiplication by the matrix A. Find
    (a) basis for the range of T.
    (b) basis for the kernel of T.
    (c) the rank and nullity of T.
    (d) the rank and nullity of A.

19. $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & -1 & 3 \\ 5 & 6 & -4 \\ 7 & 4 & 2 \end{bmatrix}$

21. $A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix}$

22. $A = \begin{bmatrix} 4 & 1 & 5 & 2 \\ 1 & 2 & 3 & 0 \end{bmatrix}$

23. $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & 1 \\ -2 & -2 & 2 \end{bmatrix}$

24. Let V be any vector space, and let $T : V \to V$ be defined by $T(\vec{v}) = 100\vec{v}$.
    (a) What is the kernel of T ?
    (b) What is the range of T ?

25. Let A be a 6×5 matrix such that $AX = 0$ has only trivial solution, and let $T : \mathbb{R}^5 \to \mathbb{R}^6$ be multiplication by A. Find rank and nullity of T.

26. Let A be 6×7 matrix with rank 4.
    (a) What is the dimension of the solution space of $AX = 0$?
    (b) Is $AX = b$ consistent for all vectors in R^6 ? Explain.

27. Let $T : V \to V$ be a linear transformation where V is finite dimensional vector space. Prove that $R(T) = V$ if and only if $\ker(T) = \{\vec{0}\}$.

Answers (5.1)

1. (a) Yes,
   (b) No, addition and scalar multiplication are not preserved.
   (c) Yes
   (d) Yes
   (e) Yes
   (f) No, addition is not preserved.
   (g) No, addition is not preserved.
   (h) No, addition is not preserved.
2. (a) Yes
   (b) No, $T(0, 0, 0) \neq (0, 0, 0)$
   (c) Yes
   (d) No, addition is not preserved
   (e) Yes
   (f) No, addition and scalar multiplication are not preserved.

3. (a) Yes, (b) No, (c) Yes, (d) No.
5. (a) For T_2oT_1 domain is R^2 and image space is $\mathbb{R}^2$ and $(T_2oT_1)\ (x, y) = (2x + 3y, x - 2y)$.
   For T_1 o T_2 domain is $\mathbb{R}^3$ and image space is $\mathbb{R}^3$.
   (b) For T_2 o T_1 domain is R^2 and image space is the line $x = 0$; $(T_2$ o $T_1)\ (x, y) = (0, 3x)$. For T_1 o T_2, domain is R^3 and the image space is the line $(t\ (-1, 1) \mid t \in \mathbb{R}\}$; that is, the line $x = y$.
7. $(17, -7, 1)$.
8. Not linear transformation.
11. $3\vec{e_1} + 4\vec{e_2} + 4\vec{e_3} = (3, 4, 4)$; nullity 0, rank 3.

    (**Hint :** Note that $\{\vec{e_3}, \vec{e_2} + \vec{e_3}, \vec{e_1} + \vec{e_2} + \vec{e_3}\}$ forms a basis for $\mathbb{R}^3$).
12. $(5, 0, -1)$; nullity 0, rank 2.
13. $(-1, -3, -1)$; nullity 0, rank 2.
14. (a) (i) , (iii), (b) (i).
15. (a) (i), (ii), (iii), (b) (i).
16. (a) $\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$, (b) $\begin{bmatrix} \frac{3}{2} \\ -4 \\ 1 \\ 0 \end{bmatrix}$
17. (a) $\begin{bmatrix} 1 \\ -4 \end{bmatrix}$, (b) $\begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -4 \\ 9 \end{bmatrix}$
19. (a) $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

    (c) rank (T) = 2, nullity (T) = 1.
20. (a) $\begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} -\frac{14}{11} \\ \frac{19}{11} \\ 1 \end{bmatrix}$

    (c) rank (T) = 1, nullity (T) = 2.
21. (a) $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

    (c) rank (T) = 1, nullity (T) = 2.

22. (a) $\begin{bmatrix}1\\ \frac{1}{4}\end{bmatrix}, \quad \begin{bmatrix}0\\ 1\end{bmatrix}$ (b) $\begin{bmatrix}-1\\ -1\\ 1\\ 0\end{bmatrix}, \quad \begin{bmatrix}-\frac{4}{7}\\ 2\\ \frac{2}{7}\\ 0\\ 1\end{bmatrix}$

(c) rank (T) = 2, nullity (T) = 2.

23. (a) $\begin{bmatrix}1\\ 1\\ -1\end{bmatrix}, \begin{bmatrix}0\\ 1\\ -\frac{2}{3}\end{bmatrix}$ (b) $\begin{bmatrix}\frac{1}{3}\\ \frac{2}{3}\\ 1\end{bmatrix}$

(c) Rank (T) = 2, nullity (T) = 1

24. Ker (T) = {0}, R(T) = V
25. Rank (T) = 5, Nullity (T) = 0
26. (a) 3, (b) No. for this rank (T) should be 6.

5.5 Inverse of a Linear Transformation

Definition : A linear transformation T : V → W is said to be non-singular if and only if it is bijective, that is, if and only if it is one-to-one and onto, otherwise, it is called singular.

Theorem 12 : Let T : V → W be an injective linear transformation and $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_k}\}$ be the set of linearly independent vectors in V. Then, the set $\{T(\vec{u_1}), T(\vec{u_2}), \ldots, T(\vec{u_k})\}$ is also linearly independent in W.

Proof : Suppose we have scalars $\alpha_1, \alpha_2, \ldots, \alpha_k$ such that

$$\alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \ldots + \alpha_k T(\vec{u_k}) = \vec{0} = T(\vec{0})$$

by property of linear transformation. As T is linear transformation of left hand side becomes

$$T(\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_k \vec{u_k}) = T(\vec{0}).$$

As T is injective, we must have

$$\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_k \vec{u_k} = \vec{0}.$$

Now, $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_k}\}$ being linearly independent, it implies that $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, thus showing that $\{T(\vec{u_1}), T(\vec{u_2}), \ldots, T\vec{u_k})\}$ is linearly independent.

Theorem 13 : A linear transformation T : V → W is non-singular if and only if a basis of V maps to the basis of W, that is, if and only if $\{T(\vec{u_1}), T(\vec{u_2}), \ldots, T(\vec{u_n})\}$ is a basis of W whenever $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ is a basis for V.

Proof : Suppose T is non-singular, so it is injective by theorem 12, if $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ is basis of V, then the set $\{T(\vec{u_1}), T(\vec{u_2}), \ldots, T(\vec{u_n})\}$ is linearly independent. Again T is surjective for any $\vec{v}$ in W, there exists a vector

$$\vec{u} = \alpha_1\vec{u_1} + \alpha_2\vec{u_2} + \ldots + \alpha_n\vec{u_n}$$

in V such that $T(\vec{u}) = \vec{v}$, that is

$$\vec{v} = \alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \ldots + \alpha_n T(\vec{u_n}),$$

Showing that $\{T(\vec{u_1}), T(\vec{u_2}), \ldots, T(\vec{u_n})\}$ spans W, and that is a basis of W.

Conversely suppose that basis of V maps to a basis of W, to show that T is non-singular. Let $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be a basis of V. Then $\{T(\vec{u_1}), T(\vec{u_2}), \ldots, T(\vec{u_n})\}$ is also a basis of W. For any two vectors $\vec{u}$ and $\vec{v}$ in V, we have

$$\vec{u} = \alpha_1\vec{u_1} + \alpha_2\vec{u_2} + \ldots + \alpha_n\vec{u_n}$$

and $$\vec{v} = \beta_1\vec{u_1} + \beta_2\vec{u_2} + \ldots + \beta_n\vec{u_n}$$

for scalars $\alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots, \beta_n$. Now suppose

$$T(\vec{u}) = T(\vec{v})$$

We have $T(\vec{u} - \vec{v}) = \vec{0}$, therefore

$$T[(\alpha_1\vec{u_1} + \alpha_2\vec{u_2} + \ldots + \alpha_n\vec{u_n}) - (\beta_1\vec{u_1} + \beta_2\vec{u_2} + \ldots + \beta_n\vec{u_n})] = \vec{0}$$

or $$T[(\alpha_1 - \beta_1)\vec{u_1} + (\alpha_2 - \beta_2)\vec{u_2} + \ldots + (\alpha_n - \beta_n)\vec{u_n}] = \vec{0}$$

Since T is linear, this gives

$$(\alpha_1 - \beta_1) T(\vec{u_1}) + (\alpha_2 - \beta_2) T(\vec{u_2}) + \ldots + (\alpha_n\ \beta_n) T(\vec{u_n}) = \vec{0}$$

This simples $\alpha_1 - \beta_1 = 0 = \alpha_2 - \beta_2 = \ldots = \alpha_n - \beta_n$.

So $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2, \ldots, \alpha_n = \beta_n$, hence $\vec{u} = \vec{v}$, and then T is injective.

Again, for any $\vec{w}$ in W, we can write

$$\vec{w} = \alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \ldots + \alpha_n T(\vec{u_n})$$

Since T is linear transformation we have

$$\vec{w} = T(\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_n \vec{u_n}) = T(\vec{u}),$$

that is, every vector of W is image of some vector in V, so T is onto. Thus T is non-singular.

We have seen the composition or product of two linear transformation in previous sections, now we define inverse lf linear transformation.

Definition : For a linear transformation $T : V \to W$ if there exists a linear transformation $S : W \to V$ such that $ST = I_v$, the identity linear transformation on V.

Then S is called the **left inverse** of T. On the other hand, if

$$W \xrightarrow{S} V \xrightarrow{T} W$$

assume that $TS = I_W$, then S is called the **right inverse** of T.

If $ST = I_V$ and $TS = I_W$ are both satisfied, then S is said to be the **inverse** of T.

Example (1) : Let T and S be linear transformations from $\mathbb{R}$ to $\mathbb{R}$, defined by $T(x) = 5x$ and $S(x) = \frac{1}{5}x$, for $x \in \mathbb{R}$. Then, we have

$$(ST)(x) = S(T(x)) = S(5x) = \frac{1}{5}(5x) = x \quad \text{and}$$

$$(TS)(x) = T(S(x)) = T\left(\frac{1}{5}x\right) = 5\left(\frac{1}{5}x\right) = x.$$

Therefore, $ST = TS = I$, T has an inverse which is S, conversely, S has inverse which is T.

SOLVED EXAMPLES

Example 5.24 : Show that the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$, defined by $T(x, y, z) = (y, x + y - z)$ is onto but not one-to-one. Find the right inverse of T.

Solution : If $(u, v) \in \mathbb{R}^2$ is any vector, then suppose $T(x, y, z) = (u, v)$, by definition T, this gives $(y, x + y - z) = (u, v)$, so that $y = u$, $x + y - z = v$ or $y = u$, $x = v - u + z$. Thus, $x = v - u + \alpha$, $y = u$, where, α is arbitrary real no, with $\alpha = z$, hence given any $(u, v) \in \mathbb{R}^2$, there are infinitely many vectors in $\mathbb{R}^3$, which map (u, v) in $\mathbb{R}^2$, hence **T is onto**. As we see that for given any $(u, v) \in \mathbb{R}^2$, there are infinitely many vectors in $\mathbb{R}^3$ which map (u, v), so T cannot be one-to-one.

Now let S be right inverse of T, so that

$T : \mathbb{R}^3 \to R^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^3$ or $\mathbb{R}^2 \xrightarrow{S} \mathbb{R}^3 \xrightarrow{T} \mathbb{R}^2$, so that $TS : \mathbb{R}^2 \to \mathbb{R}^2$.

Since S is right inverse of T, we must have

$$(TS)(u, v) = (u, v) \qquad \ldots (i)$$

Suppose we define $S(u, v) = (x, y, z)$, then

$$T(S(u, v)) = T(x, y, z) = (y, x + y - z)$$

From equation (i) this shows that

$$u = y \text{ and } v = x + y - z$$

This solving for x, y, z gives

$$y = u, \quad x = v - u + z \text{ or } y = u, \; x = v - u + \alpha, \; z = \alpha \text{ arbitrary.}$$

Thus, $S(u, v) = (v - u + \alpha, u, \alpha)$

For linearity of S, $\alpha = 0$

$\therefore \quad S(u, v) = (v - u, u, 0)$

Theorem 14 : A linear transformation $T : V \to W$ has an inverse if and only if it is bijective.

Proof : Suppose $S : W \to V$ is an inverse of T, so that $ST = I_V$ and $TS = I_W$. … (*)

Let $\vec{u}$ and $\vec{v}$ be any two vectors in V such that $T(\vec{u}) = T(\vec{v})$.

Then $(ST)(\vec{u}) = S(T(\vec{u})) = S(T(\vec{v})) = (ST)(\vec{v})$

This form (*), gives that $\vec{u} = \vec{v}$, hence L is injective. On the other hand let $\vec{w}$ be any vector in W so that $S(\vec{w}) = \vec{u}$ in V. From (*) again we have

$$(TS)(\vec{w}) = \vec{w} \quad \because \text{ TS is identity on W.}$$

Therefore, $T(\vec{u}) = \vec{w}$, where $\vec{u} = S(\vec{w})$ in V.

This means that every vector $\vec{w}$ in W is image of some vector $\vec{u}$ in V under T. Therefore, T is surjective.

Conversely, suppose T is bijective, that is, T is both injective and surjective.

As T is surjective, for every vector $\vec{w}$ in W, there is a vector $\vec{u} \in V$ such that $T(\vec{u}) = \vec{w}$. Because T is injective this vector $\vec{u}$ is uniquely determined by $\vec{w}$. This defines a correspondence from W to V, which assigns to each $\vec{w} \in W$ to $\vec{u} \in V$, and can be defined by the transformation $S(\vec{w}) = \vec{u}$. Therefore,

$$\vec{u} = S(\vec{w}) = S(T(\vec{u}) = (ST)(\vec{u}), \text{ so that } (ST)(\vec{u}) = \vec{u} \text{ and also}$$

$$\vec{w} = T(\vec{u}) = T(S(\vec{w})) = (Ts)(\vec{w}) \text{ so that } (TS)(\vec{w}) = \vec{w}, \; \vec{w} \in W.$$

This shows that S is an inverse of T.

Corollary : A linear transformation $T : V \to W$ is invertible if and only if it is non-singular.

Proof : We know that a linear transformation is non-singular iff it is bijective and proof follows from theorem (14).

Theorem 15 : A linear transformation $T : V \to W$ has

(a) a left inverse if and only if it is injective and

(b) a right inverse if and only if it is surjective.

Proof : (a) Suppose T has left inverse. Now by theorem (14). T must be injective, because a linear transformation has inverse then it has left inverse as well as right inverse. So T is injective follows from theorem (14), Conversely, suppose T is injective, then for each $\vec{w}$ in W, there is atmost one $\vec{u}$ in V such that $\vec{w} = T(\vec{u})$. Now, if $\vec{u_1}$ is an arbitrary vector in V, then the function $S : W \to V$ defined by

$$S(\vec{w}) = \vec{u} \text{ if } \vec{w} \in T(V) \text{ with } T(\vec{u}) = \vec{w}$$

$$= \vec{u_1}, \quad \text{otherwise}$$

is such that $(ST)(\vec{u}) = S(T(\vec{u})) = S(\vec{w}) = \vec{u}$, showing that $(ST)(\vec{u}) = \vec{u}$, for every $\vec{u} \in V$, implying that $ST = I_V$ and that S is the left inverse of T.

Note : That T has many left, inverses, one for each choice of $\vec{u}$.

(b) If T has right inverse, then again by theorem (14) T is surjective.

Conversely, suppose T is surjective, then to each $\vec{w} \in W$ there is at least one $\vec{u}$ in V such that $T(\vec{u}) = \vec{w}$.

Choose one such $\vec{u}$ for each $\vec{w}$ in W and let us define

$$S : W \to V$$

by setting, $S(\vec{w}) = \vec{u}$. Then

$(TS)(\vec{w}) = T(S(\vec{w})) = T(\vec{u}) = \vec{w}$ showing that $TS = I_W$. Since $T(\vec{u}) = \vec{w}$ may hold for more than one value of $\vec{u}$, there exist infinitely many right inverses of T.

Theorem 16 : A linear transformation $T : V \to W$ is injectiev if and only if $\ker T = \{\vec{0}\}$.

Proof : Suppose T is injective and for $\vec{u} \in V$, $T(\vec{u}) = \vec{0}$. Then we know that $T(\vec{0}) = \vec{0}$, we have

$$T(\vec{u}) = T(\vec{0}),$$

then T being injective this gives $\vec{u} = \vec{0}$, showing that only one vector in V maps to $\vec{0}$ in W, that is ker T = $\{\vec{0}\}$.

Conversely, suppose ker T = $\{\vec{0}\}$. If $\vec{u}$ and $\vec{v}$ are vectors in V, and suppose

$$T(\vec{u}) = T(\vec{v})$$

This implies $T(\vec{u}) - T(\vec{v}) = \vec{0}$, since T is linear, this gives $T(\vec{u} - \vec{v}) = \vec{0}$, so that $\vec{u} - \vec{v} \in \ker T$. But ker T = $\{\vec{0}\}$, hence we must have $\vec{u} - \vec{v} = \vec{0}$ or $\vec{u} = \vec{v}$. This shows that T is injective.

SOLVED EXAMPLES

Example 5.25 : Find a linear transformation T : $\mathbb{R}^3 \to \mathbb{R}^2$ which maps the vectors $\vec{u_1} = (1, 1, 1)$, $\vec{u_2} = (1, 1, 0)$, $\vec{u_3} = (1, 0, 0)$ in $\mathbb{R}^3$ to the vectors $\vec{v_1} = (2, 1)$, $\vec{v_2} = (2, 1)$, $\vec{v}_3 = (2, 1)$ respectively.

Solution : The vectors $\vec{u_1}, \vec{u_2}, \vec{u_3}$ form a basis for $\mathbb{R}^3$. So each vector in $\mathbb{R}^3$ is a linear combination of $\vec{u_1}, \vec{u_2}, \vec{u_3}$. If $\vec{u} = (a, b, c)$ is any vector in $\mathbb{R}^3$, we have scalars α, β, δ such that

$$\begin{aligned} \vec{u} = (a, b, c) &= \alpha\vec{u_1} + \beta\vec{u_2} + \delta\vec{u_3} \qquad \ldots \text{(i)} \\ &= \alpha(1, 1, 1) + \beta(1, 1, 0) + \delta(1, 0, 0) \\ &= (\alpha + \beta v + \delta, \alpha + \beta, \alpha) \end{aligned}$$

So equating the corresponding components, we get

$$a = \alpha + \beta + \delta, \; b = \alpha + \beta \text{ and } c = \alpha$$

solving for α, β, δ, this gives

$$\alpha = c, \quad \beta = b - c \quad \text{and} \quad \delta = a - b$$

Thus, from equation (i) above, we have

$$\vec{u} = (a, b, c) = c(1, 1, 1) + (b - c)(1, 1, 0) + (a - b)(1, 0, 0)$$

Applying T and using that T is linear we obtain

$$\begin{aligned} T(a, b, c) &= c\,T(1, 1, 1) + (b - c)\,T(1, 1, 0) + (a - b)\,T(1, 0, 0) \\ &= c(2, 1) + (b - c)(2, 1) + (a - b)(2, 1) \\ &= (2c + 2b - 2c + 2a - 2b, c + b - c + a - b) \\ &= (2a, a) \end{aligned}$$

Or $T(x, y, z) = (2x, x)$ for $(x, y, z) \in \mathbb{R}^3$, a formula for T.

Example 5.26 : A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$T(x, y, z) = (-x - 2y + z,\ x + y,\ x)$$

Find another linear transformation S such that $(ST)(\vec{u}) = (TS)(\vec{u}) = \vec{u}$, for each $\vec{u} \in \mathbb{R}^3$, that is find inverse transformation S of T.

Solution : First we show that such a linear transformation S exists by showing that T is bijective.

If $(x, y, z) \in \ker T$, then $T(x, y, z) = (0, 0, 0)$, from definition of T, this gives that $-x - 2y + z = 0$, $x + y = 0$, $x = 0$. Solving these equations, we get $x = 0$, $y = 0$, $z = 0$, showing that $\ker T = \{(0, 0, 0)\}$. Therefore, by theorem (16) **T is injective map**.

Now, to see that T is surjective, suppose $\vec{w} = (a, b, c)$ be any vector in $\mathbb{R}^3$ (codomain). Suppose

$$(a, b, c) = T(x, y, z) \quad \ldots (i)$$

We will show that there really exists $(x, y, z) \in \mathbb{R}^3$ (domain) satisfying (i). By definition of T, we get

$$(a, b, c) = (-x - 2y + z,\ x + y,\ x)$$

On equating corresponding components, we get

$$a = -x - 2y + z,\ b = x + y,\ c = x$$

Solving for x, y, z this gives

$$x = c,\ y = b - c \text{ and } z = a + 2b - c$$

This shows that every vector (a, b, c) in $\mathbb{R}^3$ (codomain) there exists a vector $(x, y, z) \in \mathbb{R}^3$ (domain) such that

$$T(x, y, z) = (a, b, c)$$

which shows that T is surjective.

Then, as T is bijective by theorem inverse of T exists. Let S be the inverse of T, then by definition

$$ST = TS = I, \text{ where } S : \mathbb{R}^3 \to \mathbb{R}^3$$

This shows that

$$(ST)(x, y, z) = (TS)(x, y, z) = (x, y, z) \quad \ldots (ii)$$

for every (x, y, z) in $\mathbb{R}^3$.

Now from definition of T and the fact that (1, 0, 0), (0, 1, 0), (0, 0, 1) form a standard basis for $\mathbb{R}^3$, we have

$$T(1, 0, 0) = (-1, 1, 1),\ T(0, 1, 0) = (-2, 1, 0),\ T(0, 0, 1) = (1, 0, 0)$$

The fact that the bijective linear transformation maps basis to basis, we see that $\vec{w_1} = (-1, 1, 1)$, $\vec{w_2} = (-2, 1, 0)$, $\vec{w_3} = (1, 0, 0)$ forms a basis for $\mathbb{R}^3$. We can define a uniquely determined linear transformation $S : \mathbb{R}^3 \to \mathbb{R}^3$ by setting

$$S(\vec{w_1}) = (1, 0, 0),\ S(\vec{w_2}) = (0, 1, 0),\ S(\vec{w_3}) = (0, 0, 1).$$

Next for any $\vec{w} = (a, b, c)$ in $\mathbb{R}^3$ (codomain),

$$\vec{w} = (a, b, c) = \alpha_1\vec{w_1} + \alpha_2\vec{w_2} + \alpha_3\vec{w_3} \quad \ldots \text{(iii)}$$

Since $\vec{w_1}, \vec{w_2}, \vec{w_3}$ form a basis for $\mathbb{R}^3$.

$$\therefore \quad \vec{w} = \alpha_1 (-1, 1, 1) + \alpha_2 (-2, 1, 0) + \alpha_3 (1, 0, 0)$$
$$= (-\alpha_1 - 2\alpha_2 + \alpha_3,\ \alpha_1 + \alpha_2,\ \alpha_1)$$

Equating corresponding components, we get

$$a = -\alpha_1 - 2\alpha_2 + \alpha_3, \quad b = \alpha_1 + \alpha_2, \quad c = \alpha_1$$

Solving these equations for α_1, α_2 and α_3 we obtain

$$\alpha_1 = c, \quad \alpha_2 = b - c, \quad \alpha_3 = a + 2b - c$$

Applying S on both sides of (iii) and as S is linear transformation, we obtain

$$S(\vec{w}) = S(a, b, c) = c\,S(-1, 1, 1) + (b - c)\,S(-2, 1, 0) + (a + 2b - c)\,S(1, 0, 0)$$
$$= c\,(1, 0, 0) + (b - c)\,(0, 1, 0) + (a + 2b - c)\,(0, 0, 1)$$
$$= (c, b - c, a + 2b - c)$$

Thus $S(a, b, c) = (c, b - c, a + 2b - c)$

The required inverse of T.

Example 5.27 : Let $D = \frac{d}{dx}$ and T be linear transformations on p(x), the vector space of polynomials defined by

$$D(p(x)) = \frac{d}{dx}(p(x)) \text{ and } T(p(x)) = x\,P(x))$$

Show that DT – TD is an identity transformation.

Solution : For any p(x) is P(x), we have

$$(DT - TD)(p(x)) = (DT)(p(x)) - (TD)(p(x))$$
$$= D[T(p(x))] - T[D(p(x))]$$
$$= D[x\,p(x)] - T\left(\frac{d}{dx}p(x)\right)$$
$$= \frac{d}{dx}[x\,p(x)] - x\frac{d}{dx}(p(x))$$
$$= p(x) + x\frac{d}{dx}(p(x)) - x\frac{d}{dx}(p(x))$$
$$= p(x)$$

Thus **DT – TD** is an identity transformation.

Example 5.28 : Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T(x, y, z) = (x + y, y + z, z + x)$. Show that T is linear isomorphism. Find similar formula for T^{-1}.

Solution : One can check that T is a linear map. We show that T is 1-to-1 and onto.

To show that T is 1-to-1, we show that ker (T) = {(0, 0, 0)}. For suppose, $(x, y, z) \in$ ker (T), then

$$T(x, y, z) = (0, 0, 0), \text{ but}$$
$$T(x, y, z) = (x + y, y + z, z + x)$$

$\therefore$ We must have, $x + y = 0$, $y + z = 0$ and $z + x = 0$.

Last two equation show that $z = -x$ and $y = x$, hence using in first we see that $x + x = 0$, so that $x = 0$ and then $y = 0$, $z = 0$. Hence T is 1-to-1.

Now to show that T is onto : Let (x_1, y_1, z_1) be any point in $\mathbb{R}^3$ (co-domain) and suppose $T(x, y, z) = (x_1, y_1, z_1)$, then we must have $(x + y, y + z, z + x) = (x_1, y_1, z_1)$. So we have three equations $x + y = x_1$, $y + z = y_1$, $z + x = z_1$.

From first and third equation, we have $x = x_1 - y$ and $x = z_1 - z$, so that

$$x_1 - y = z_1 - z \quad \text{or} \quad y - z = x_1 - z_1$$

Second equation and the last equation by adding we get $y = \frac{1}{2}(x_1 + y_1 - z_1)$. Using this in first equation we get

$$x = \frac{1}{2}(x_1 + z_1 - y_1)$$

Similarly, we find $z = \frac{1}{2}(z_1 + y_1 - x_1)$.

Clearly $x = \frac{1}{2}(x_1 + z_1 - y_1)$,

$$y = \frac{1}{2}(x_1 + y_1 - z_1),\ z = \frac{1}{2}(y_1 + z_1 - x_1)$$

So that $(x, y, z) \in \mathbb{R}^3$ (domain) for any $(x_1, y_1, z_1) \in \mathbb{R}^3$ (domain) for any $(x_1, y_1, z_1) \in \mathbb{R}^3$ (codomain), such that $T(x, y, z) = (x_1, y_1, z_1)$ hence T is onto. Thus T is isomorphism. Now, to find formula for T^{-1}, we have from above computations.

$$T^{-1}(x_1, y_1, z_1) = (x, y, z)$$

$\therefore$ $$T^{-1}(x_1, y_1, z_1) = \left(\frac{1}{2}(x_1 + z_1 - y_1), \frac{1}{2}(x_1 + y_1 - z_1), \frac{1}{2}(y_1 + z_1 - x_1)\right)$$

which is a formula for T^{-1}.

Example 5.29 : Let $B = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ be basis for $\mathbb{R}^3$, where $\vec{u_1} = (-1, 0, 1)$, $\vec{u_2} = (0, 1, -1)$, $\vec{u_3} = (1, -1, 1)$. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be linear transformation for which $T(\vec{u_1}) = (1, 0, 0) = \vec{e_1}$, $T(\vec{u_2}) = (0, 1, 0) = \vec{e_2}$ and $T(\vec{u_3}) = (0, 0, 1) = \vec{e_3}$, $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ is a standard basis for $\mathbb{R}^3$. Find formula T and use it find T (2, 1, –3).

Solution : For any (x, y, z) in $\mathbb{R}^3$, suppose we have

$$\begin{aligned}(x, y, z) &= \alpha\vec{u_1} + \beta\vec{u_2} + \delta\vec{u_3}\\ &= \alpha(-1, 0, 1) + \beta(0, 1, -1) + \delta(1, -1, 1)\\ &= (-\alpha + \delta, \beta - \delta, \alpha - \beta + \delta)\end{aligned}$$

This gives

$$\begin{aligned}-\alpha + \delta &= x\\ \beta - \delta &= y\\ \alpha - \beta + \delta &= z\end{aligned} \quad \text{or} \quad \begin{aligned}\alpha - \beta + \delta &= z\\ \beta - \delta &= y\\ -\beta + 2\delta &= x + z\end{aligned}$$

Last two equation yield that $\delta = x + y + z$ and then $\beta = x + 2y + z$ using these two in first equation we obtain $\alpha = y + z$.

Thus, $\alpha = y + z$, $\beta = x + 2y + z$, $\delta = x + y + z$

Now, $T(x, y, z) = \alpha T(\vec{u_1}) + \beta T(\vec{u_2}) + \delta(\vec{u_3})$, using linearity of T

$= (y + z)(1, 0, 0) + (x + 2y + z)(0, 1, 0) + (x + y + z)(0, 0, 1)$

$= (y + z, x + 2y + z, x + y + z)$

Thus, $T(x, y, z) = (y + z, x + 2y + 2, x + y + z)$

which is a formula for T.

Now, $T(2, 1, -3) = (1 - 3, 2 + 2 - 3, 2 + 1 - 3)$

$= (-2, 1, 0)$

5.6 Matrix of a Linear Transformation

In this section we will see the explicit computations with linear transformations can be translated into computations with matrices having scalars as entries.

Let $T : V \to W$ be a linear transformation from a vector space V to the vector space W over field **F**. Let $B = \{\vec{u_1}, \ldots, \vec{u_n}\}$ be a basis of V and $B' = \{\vec{v_1}, \ldots, \vec{v_m}\}$ be the basis that of W. Since every vector $T(\vec{u_i})$, i = 1, 2, …, n in W, hence it is a linear combination of vector in B', say

$$T(\vec{u_i}) = a_{1i}\vec{v_1} + a_{2i}\vec{v_2} + \ldots + a_{mi}\vec{v_m} \qquad \ldots (i)$$

where the scalars $a_{ij's}$ are uniquely determined by the basis B'. Now, if $\vec{u} \in V$ is any vector, it is linear combination of vectors in B, let

$$\vec{u} = \alpha_1\vec{u_1} + \alpha_2\vec{u_2} + \ldots + \alpha_n\vec{u_n},$$

where the scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ are uniquely determined by a basis B of V. Suppose $\vec{u}$ is transformed to $\vec{w}$ in W under T, so that

$$\vec{w} = T(\vec{u}) = T(\alpha_1\vec{u_1} + \alpha_2\vec{u_2} + \ldots + \alpha_n\vec{u_n})$$

$$= \alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \ldots + \alpha_n T(\vec{u_n})$$

$$= \alpha_1 (a_{11} \vec{v_1} + a_{21} \vec{v_2} + \ldots + a_{m1} \vec{v_m}) + \ldots$$

$$+ \alpha_n (a_{1n} \vec{v_1} + a_{2n} \vec{v_2} + \ldots + a_{mn} \vec{v_m}) \quad \ldots \text{using (i)}$$

$$= (\alpha_1 a_{11} + \alpha_2 a_{12} + \ldots + \alpha_n a_{1n}) \vec{v_1} + \ldots$$

$$+ (\alpha_1 a_{m1} + \alpha_2 a_{m2} + \ldots + \alpha_n a_{mn}) \vec{v_m}$$

Thus, $\vec{w} = T(\vec{u}) = \beta_1 \vec{v_1} + \ldots + \beta_m \vec{v_m}$ … (ii)

where, $\beta_i = \alpha_1 a_{i1} + \alpha_2 a_{i2} + \ldots + \alpha_n a_{in}$ … (iii)

for i = 1, 2, …, m.

From equation (ii) we see that the computations with linear transformation T:V → W as far as assigning of a vector $\vec{v}$ in V to a vector $\vec{v} = T(\vec{u})$ in W is concerned can be done if we know the co-ordinate vector $(\beta_1, \beta_2, \ldots, \beta_m)$ of the vector $\vec{v}$ with respect to B', are known.

The equations (iii) can be written explicitly in matrix form as

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ : \\ \beta_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ : & : & & : \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ : \\ \alpha_n \end{bmatrix} \quad \ldots \text{(iv)}$$

The equation (iv) shows that the determination of the co-ordinate vector $(\beta_1, \beta_2, \ldots, \beta_m)$ of $\vec{v}$ can be found if we know the matrix $A = [a_{ij}]$ and the co-ordinate vector $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of $\vec{u}$ with respect to basis B. Given a linear transformation T, the scalars a_{ij} in the matrix $A = [a_{ij}]$ are uniquely determined by the relation (i) with respect to the basis B and B' so are co-ordinates of the given vector $T(\vec{u_i})$.

Definition : Let T : V → W be a linear transformation, let $B = \{\vec{u_1}, \ldots, \vec{u_n}\}$ be a basis of V and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}\}$ be the basis that of W. The matrix $A = [a_{ij}]$ uniquely determined by the relation (i) is called the **matrix representation of T** or **matrix of T** with respect to the given bases B and B' that of V and W respectively.

Note :

(1) Given a linear transformation T : V → W, the matrix of T is unique for given bases B and B' of V and W respectively. If the bases are changed the matrix of T will change.

(2) If T : V → V is an operator, the bases B and B' are taken same.

SOLVED EXAMPLES

Example 5.30 : Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation given by

$$T(x_1, x_2, x_3) = (-x_1 - x_2 + x_3, x_1 - 4x_2 + x_3, 2x_1 - 5x_2),$$

then find the matrix with respect to the standard basis $B = \{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$, where $\vec{e_1} = (1, 0, 0)$, $\vec{e_2} = (0, 1, 0)$ and $\vec{e_3} = (0, 0, 1)$.

Solution : Here we have, by definition of T,

$$T(\vec{e_1}) = T(1, 0, 0) = (-1, 1, 2)$$

$$T(\vec{e_2}) = T(0, 1, 0) = (-1, -4, -5) \text{ and}$$

$$T(\vec{e_3}) = T(0, 0, 1) = (1, 1, 0)$$

Now, we can write $T(\vec{e_1})$, $T(\vec{e_2})$ and $T(\vec{e_3})$ as a linear combination of $\vec{e_1}$, $\vec{e_2}$ and $\vec{e_3}$. It is clear that

$$T(\vec{e_1}) = (-1, 1, 2) = (-1)\,\vec{e_1} + 1\vec{e_2} + 2\vec{e_3}$$

$$T(\vec{e_2}) = (-1, -4, -5) = (-1)\,\vec{e_1} + (-4)\,\vec{e_2} + (-5)\,\vec{e_3}$$

and $$T(\vec{e_3}) = (1, 1, 0) = 1\vec{e_1} + 1\vec{e_2} + 0\cdot\vec{e_3}$$

Thus, the matrix of T with respect B is

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & -4 & 1 \\ 2 & -5 & 0 \end{bmatrix}$$

Example 5.31 : Consider the vector space $\mathbb{C}^2$ over the real field $\mathbb{R}$. Show that

$B = \{\vec{e_1}, \vec{e_2}, \vec{e_3}, \vec{e_4}\}$ is a basis of $\mathbb{C}^2$, where $\vec{e_1} = (1, 0)$, $\vec{e_2} = (i, 0)$, $\vec{e_3} = (0, 1)$ and $\vec{e_4} = (0, i)$.

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be the linear transformation given by

$$T(a + ib, c + id) = (a - ib, c - id)$$

over $\mathbb{R}$. Find the matrix representation of T with respect to the basis B.

Solution : Any vector $\vec{u}$ in $\mathbb{C}^2$ is given by

$$\vec{u} = (a + ib, c + id) \text{ with } a, b, c, d \in \mathbb{R}.$$

We can see that $\vec{u} = a(1, 0) + b(i, 0) = c(0, 1) + d(0, i)$

Showing that B spans $\mathbb{C}^2$ over $\mathbb{R}$.

To see that B is linearly independent, suppose

$$\alpha_1\vec{e_1} + \alpha_2\vec{e_2} + \alpha_3\vec{e_3} + \alpha_4\vec{e_4} = (0, 0)$$

This gives

$$(\alpha_1 + i\alpha_2, \alpha_3 + i\alpha_4) = (0, 0)$$

$\Leftrightarrow \alpha_1 + i\alpha_2 = 0$ and $\alpha_3 + i\alpha_4 = 0$

$\Leftrightarrow \alpha_1 = 0, \quad \alpha_2 = 0$ and $\alpha_3 = 0, \ \alpha_4 = 0.$

$\therefore$ B is linearly independent and hence is a basis of $\mathbb{C}^2$ over $\mathbb{R}$.

Now,
$$T(\vec{e_1}) = T(1 + 0i, 0 + 0i) = (1, 0) = 1\vec{e_1} + 0\vec{e_2} + 0\vec{e_3} + 0\vec{e_4}$$
$$T(\vec{e_2}) = T(0 + i, 0 + 0i) = (-i, 0) = 0\vec{e_1} + (-1)\vec{e_1} + 0\vec{e_3} + 0\vec{e_4}$$
$$T(\vec{e_3}) = T(0 + 0i, 1 + 0i) = (0, 1) = 0\vec{e_1} + 0\vec{e_2} + 1\vec{e_3} + 0\vec{e_4}$$
$$T(\vec{e_4}) = T(0, i) = (0, -i) = 0\cdot\vec{e_1} + 0\cdot\vec{e_2} + 0\vec{e_3} + (-1)\vec{e_4}$$

Therefore the matrix of T with respect to B is

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

From relations (i), (ii), (iii) and (iv) of the discussion in the beginning of this section, we **observe :** Given a linear transformation $T : V \to W$ and the bases B and B' of V and W respectively and a vector $\vec{u}$ in V, we can compute the vector $\vec{v} = T(\vec{u})$ in W by first determining the matrix $A = [a_{ij}]$ whose elements a_{ij}'s are uniquely determined by the relation (i) with respect B and B' and the co-ordinate vector $\vec{x} = (\alpha, \alpha_2, \ldots, \alpha_n)$ of $\vec{u}$ with respect to B. With the help of A and $\vec{x}$ so found using relation (iv), we can find the co-ordinate vector $\vec{y} = (\beta_1, \beta_2, \ldots, \beta_m)$ of $\vec{v}$ with respect B'. Then putting this in (ii) the relation $\vec{v} = T(\vec{u})$ is specified. Thus, the computation of the linear transformation

$T(\vec{u}) = \vec{v}$ is closely associated with the matrix relation $A\vec{x} = \vec{y}$, where $\vec{x}$ and $\vec{y}$ are considered as column matrices.

Conversely, given any arbitrary $m \times n$ matrix of scalar entries, the bases B and B' of vector spaces V and W respectively, there is a unique linear transformation $T : V \to W$, whose action on $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}$ is defined by the relation (i).

Thus, we have a theorem pertaining to the following discussions :

Theorem 17 : Let $T : V \to W$ be a linear transformation from a vector space V to the vector space W with bases $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}\}$ of V and W respectively.

If $\vec{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is the co-ordinate vector of $\vec{u}$ with respect to B, and $\vec{y} = (\beta_1, \beta_2, \ldots, \beta_n)$ that of a vector $\vec{v}$ in W, and a matrix $A = [a_{ij}]$ of linear transformation T with respect to the bases B and B', then $T(\vec{u}) = \vec{v}$ if and only if $A\vec{x} = \vec{y}$.

Proof : Suppose $A\vec{x} = \vec{y}$. By definition,

$\vec{u} = \alpha_1\vec{u_1} + \alpha_2\vec{u_2} \ldots + \alpha_n\vec{u_n}$ and $\vec{v} = \beta_1\vec{v_1} + \beta_2\vec{v_2} + \ldots + \beta_m\vec{v_m}$.

Using second relation in (ii), and retracing the steps we get $T(\vec{u}) = \vec{v}$.

Conversely, if we assume $T(\vec{u}) = \vec{v}$. It follows from relation (iii) that $A\vec{x} = \vec{y}$. Thus the vector equation $T(\vec{u}) = \vec{v}$ corresponds to the matrix equation $A\vec{x} = \vec{y}$ and vice versa.

Corollary : Let $T : V \to W$ be a linear transformation whose matrix representation with respect to the bases $B = \{\vec{u_1}, \ldots, \vec{u_n}\}$ and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}\}$ of V and W respectively is $A = [a_{ij}]$. Prove that T is invertible if and only if A is invertible.

Proof : Let $\vec{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $\vec{y} = (\beta_1, \beta_2, \ldots, \beta_m)$ be the co-ordinate vectors of $\vec{u}$ and $\vec{v}$ respectively, with respect to the given bases.

Suppose T is invertible then T is bijective. First note that if $T(\vec{u}) = \vec{v} = \vec{0}$ and $\vec{u} = 0$, then $\vec{x} = 0$ and $\vec{y} = 0$, because the co-ordinate vector of zero vector is zero. Now T is bijective, so it is one-to-one, hence $T(\vec{u}) = \vec{0}$ implies $\vec{u} = \vec{0}$, then from theorem

(17) $A\vec{x} = \vec{0}$ implies that $\vec{x} = \vec{0}$, showing that A is also one-to-one. Next T is onto says that for every $\vec{v}$ in W there exists $\vec{u}$ in V such that $T(\vec{u}) = \vec{v}$. Again by theorem (17) we get for every $\vec{y}$ there exists $\vec{x}$ such that $A\vec{x} = \vec{y}$. Hence A is onto and therefore A is invertible.

Conversely, if we assume that A is invertible, we can retrace the steps in above proof and show that T is invertible.

Matrix of an Identity and a Zero Linear Transformation :

The matrix of an identity linear transformation $I_v : V \to V$ with respect to any basis $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ of V, using relation (i), we have $T(\vec{u_i}) = 1\vec{u_i}$, i = 1, 2, …, n which shows that $a_{ij} = 0$ if $i \neq j$ and $a_{ij} = 1$ in (i). Thus the matrix of I_v is I_n, the unit matrix of order n, where n = dim V. So irrespective of the bases, the matrix of an identity transformation is a **unit** matrix.

Similarly, we can see that the zero transformation $\bar{0} : V \to W$, which maps every vector in V to a zero vector in W, hence the matrix of $\bar{0}$ is a zero matrix of order $m \times n$, which is not dependent on the bases of V and W.

Matrix of the sum of Two Linear Transformations and a Scalar Multiple of a Linear Transformation :

Let T_1 and T_2 be two linear transformations from a vector space V to the vector space W. Let $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ be a basis of V and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}\}$ that of W. Then from definition of addition, the relation (i), takes the forms

$$\begin{aligned}(T_1 + T_2)(u_i) &= T_1(\vec{u_i}) + T_2(\vec{u_i}),\ i = 1, 2, \ldots, n \\ &= (a_{1i}\vec{v_1} + \ldots + a_{mi}\vec{v_m}) + (b_{1i}\vec{v_1} + \ldots + b_{mi}\vec{v_m})\end{aligned}$$

and for any scalar α,

$$\begin{aligned}(\alpha T_1)(\vec{u_i}) &= \alpha\, T(\vec{u_i}),\ i = 1, 2, \ldots, n \\ &= \alpha\,(a_{1i}\vec{v_1} + a_{2i}\vec{v_2} + \ldots + a_{mi}\vec{v_m}) \\ &= (\alpha a_{1i})\vec{v_1} + (\alpha a_{2i})\vec{v_2} + \ldots + (\alpha a_{mi})\vec{v_m},\ i = 1, 2, \ldots, n.\end{aligned}$$

where a_{ij}'s and b_{ij}'s are uniquely determined scalars w.r.t. basis B' and determine the matrices $A_1 = [a_{ij}]$ and $A_2 = [b_{ij}]$. Hence, the equations corresponding to (iii) for $T_1 + T_2$ and αT_1 now becomes

$$\beta_i = \alpha_1 (a_{i1} + b_{i1}) + \alpha_2 (a_{i2} + b_{i2}) + \ldots + \alpha_n (a_{in} + b_{in}) \qquad i = 1, 2, \ldots, m$$

and $$\beta_i = \alpha_1 (\alpha a_{i1}) + \alpha_2 (\alpha a_{i2}) + \ldots + \alpha_n (\alpha a_{in}),\ i = 1, 2, \ldots, m$$

respectively. From this it follows that the matrix of $T_1 + T_2$ is the sum $A_1 + A_2$ of their individual matrices with respect B and B' and the matrix of αT_1 is α-times the matrix of T_1.

The Matrix of a Composite Linear Transformation :

Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations, where U, V and W are vector spaces over the field F. Let

$$B_1 = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\},\ B_2 = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \text{ and}$$

$$B_3 = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m\} \text{ be bases of U, V and W respectively.}$$

We denote by $A = [a_{ki}]_{n \times p}$ and $B = [b_{ik}]_{m \times n}$ the matrices of T_1 and T_2. The scalars a_{ki}'s are determined with respect to the bases B_1 and B_2, where as the scalars b_{ik}'s are determined with respect to the bases B_2 and B_3.

Then we have,

$$T_1(\vec{u}_j) = a_{1j}\vec{v}_1 + \ldots + a_{kj}\vec{v}_k + \ldots + a_{nj}\vec{v}_n$$

and
$$T_2(\vec{v}_k) = b_{1k}\vec{w}_1 + \ldots + b_{ik}\vec{w}_k + \ldots + b_{mk}\vec{w}_m$$

Now, using the linearity of T_1 and T_2, we have

$$(T_2 T_1)(\vec{u}_j) = T_2[T_1(\vec{u}_j)] = \sum_{k=1}^{n} a_{kj} T_2(\vec{v}_k)$$

$$= \sum_{k=1}^{n}\sum_{i=1}^{m} a_{kj} b_{ik}\vec{w}_i = \sum_{i=1}^{m} c_{ij}\vec{w}_i$$

where,
$$c_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}$$

which shows that the matrix of the composite transformation T_2T_1 is $C = [c_{ij}] = BA$. In particular matrix of T_1^2 will be A^2 and that T_1^r, r positive integer, will be A^r.

Matrix of an Inverse Linear Transformation :

Let $T : V \to W$ be a linear transformation and $S : W \to V$ be the inverse of T. It is clear from the fact that $ST = I_v$ and $TS = I_w$ and T is bijective, dim V = dim W. Let

$B = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ and $B' = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be bases of V and W respectively.

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ are matrix representation of T and S respectively with respect to these bases then we have, in view of relation (i), the relation :

$$T(\vec{u}_k) = a_{1k}\vec{v}_1 + \ldots + a_{ik}\vec{v}_i + \ldots + a_{nk}\vec{v}_n$$

and
$$S(\vec{v}_j) = b_{1j}\vec{u}_1 + \ldots + b_{kj}\vec{u}_k + \ldots + b_{nj}\vec{u}_n,\ k, j = 1 \text{ to } n.$$

Now, since T is linear,

$$(TS)(\vec{v_j}) = T[S(\vec{v_j})] = \sum_{k=1}^{n} b_{kj}\, T(\vec{u_k})$$

$$= \sum_{k=1}^{n} b_{kj} \sum_{i=1}^{n} a_{ik}\, \vec{v_i} = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik}\, b_{kj} \right) \vec{v_i}$$

So that the matrix of TS with respect to given basis B' is AB. Again T is right invertible, $TS = I_W$, which shows that the matrix of TS is alternatively given by I_n with respect to the same basis B'. Since the matrix of a linear transformation is unique with respect to given basis, we must have $AB = I_n$. Next, when T is considered to be left invertible, we get $BA = I_n$. So the matrix A of the linear transformation T has an inverse equal to the matrix B of its inverse transformation S.

In example (8) and (9) we obtained matrix of given linear transformation almost straight forward because the bases used were standard. However, it requires good many calculations, like solving systems of linear equation, then in some examples we need more efforts.

SOLVED EXAMPLES

Example 5.32 : Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation given by $T(x, y, z) = (-x - y + z,\ x - 4y + z,\ 2x - 5y)$, determine the matrix of T with respect to the basis $B = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$, where $\vec{u_1} = (1, 0, 2)$, $\vec{u_2} = (2, 1, 0)$, $\vec{u_3} = (1, 0, 1)$.

Solution : We first find $T(\vec{u_1})$, $T(\vec{u_2})$ and $T(\vec{u_3})$ from definition of T and then we express these three vectors as linear combination of $\vec{u_1}$, $\vec{u_2}$ and $\vec{u_3}$. So

$$T(\vec{u_1}) = T(1, 0, 2) = (1, 3, 2)$$

$$T(\vec{u_2}) = T(2, 1, 0) = (-3, -2, -3)$$

$$T(\vec{u_3}) = (0, 2, 2)$$

Now suppose, $(1, 3, 2) = \alpha\vec{u_1} + \beta\vec{u_2} + \delta\vec{u_3}$

$$= \alpha(1, 0, 2) + \beta(2, 1, 0) + \delta(1, 0, 1)$$

$$= (\alpha + 2\beta + \delta,\ \beta,\ 2\alpha + \delta)$$

Equating the corresponding components, we get

$$\alpha + 2\beta + \delta = 1, \quad \beta = 3, \quad 2\alpha + \delta = 2$$

Solving this system for α, β and δ, we obtain

$$\alpha = 7, \quad \beta = 3, \quad \delta = -12$$

Thus, $T(\vec{u_1}) = (1, 3, 2) = 7\vec{u_1} + 3\vec{u_2} + (-12)\,\vec{u_3}$ … (i)

Similarly, writing $T(\vec{u_2}) = \alpha\vec{u_1} + \beta\vec{u_2} + \delta\vec{u_3} = (-3, -2, -3)$ gives

$$\alpha = -4, \quad \beta = -2, \quad \delta = 5,$$

hence $$T(\vec{u_2}) = (-3, -2, -3) = (-4)\,\vec{u_1} + (-2)\,\vec{u_2} + 5\vec{u_3} \quad \ldots \text{(ii)}$$

Also $$T(\vec{u_3}) = (0, 2, 2) = 6\vec{u_1} + 2\vec{u_2} + (-10)\,\vec{u_3} \quad \ldots \text{(iii)}$$

Thus, the matrix of T w.r.t. the basis B is

$$A = [a_{ij}] = \begin{bmatrix} 7 & -4 & 6 \\ 3 & -2 & 2 \\ -12 & 5 & -10 \end{bmatrix}$$

Example 5.33 : Find the matrix representation of the derivative operator $D: P_4(x) \to P_3(x)$ with respect to the bases

$$B = \{1, x, x^2, x^3, x^4\} \text{ and } B' = \{1, 1 + x, x + x^2, x^2 + x^3\}.$$

Solution : We have, by definition of D, $D = \dfrac{d}{dx}$

$$\begin{aligned} D(1) = 0 &= 0 \cdot 1 + 0 \cdot (1 + x) + 0 \cdot (x + x^2) + 0\,(x^2 + x^3) \\ D(x) = 1 &= 1 \cdot 1 + 0 \cdot (1 + x) + 0 \cdot (x + x^2) + 0 \cdot (x^2 + x^3) \\ D(x^2) = 2x &= -2 + 2\,(1 + x) + 0 \cdot (x + x^2) + 0 \cdot (x^2 + x^3) \\ D(x^3) = 3x^2 &= 3 + (-3)\,(1 + x) + 3\,(x + x^2) + 0 \cdot (x^2 + x^3) \end{aligned}$$

and $$D(x^4) = 4x^3 = -4 + 4(1 + x) + (-4)\,(x + x^2) + 4\,(x^2 + x^3)$$

Thus, the matrix of D is $A = \begin{bmatrix} 0 & 1 & -2 & 3 & -4 \\ 0 & 0 & 2 & -3 & 4 \\ 0 & 0 & 0 & 3 & -4 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}_{4 \times 5}$

For students as an exercise, find the matrix of $D : P_n(x) \to P_{n-1}(x)$ with respect to the bases $B = \{1, x, x^2, \ldots, x^n\}$ and $B' = \{1, 1 + x, x + x^2, \ldots, x^{n-2} + x^{n-1}\}$. This generalization of the above example (11).

Example 5.34 : Find the matrix of the sum of two linear transformations $T_1: \mathbb{R}^2 \to \mathbb{R}^3$ and $T_2 : \mathbb{R}^2 \to \mathbb{R}^3$ given by $T_1(x, y) = (x + y, x, x - y)$ and $T_2(x, y) = (x + 2y, x - y, y)$ with respect to the bases $B = \{(1, 0), (1, 1)\}$ and $B' = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ of $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively.

Solution : Let us write

$\vec{u_1} = (1, 0)$, $\vec{u_2} = (1, 1)$ and $\vec{v_1} = (1, 0, 0)$, $\vec{v_2} = (1, 1, 0)$, $\vec{v_3} = (1, 1, 1)$.

First we find matrix of T_1. We have,

$$\begin{aligned} T(\vec{u_1}) &= T(1, 0) = (1, 1, 1) = 0 \cdot \vec{v_1} + 0 \cdot \vec{v_2} + 1 \cdot \vec{v_3} \\ T(\vec{u_2}) &= T(1, 1) = (2, 1, 0) = 1 \cdot \vec{v_1} + 1 \cdot \vec{v_2} + 0 \cdot \vec{v_3} \end{aligned}$$

Thus, the matrix of T_1 is

$$A = [a_{ij}] = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Next, for matrix of T_2,

$$T_2(\vec{u_1}) = T(1, 0) = (1, 1, 0) = 0 \cdot \vec{v_1} + 1 \cdot \vec{v_2} + 0 \cdot \vec{v_3}$$

and $$T_2(\vec{u_2}) = T(1, 1) = (3, 0, 1) = 3\vec{v_1} + (-1)\vec{v_2} + 1 \cdot \vec{v_3}$$

Then, the matrix of T_2 is

$$B = [b_{ij}] = \begin{bmatrix} 0 & 3 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Now, we have seen that matrix of the sum of two linear transformations is the sum of their respective matrix representations. Therefore, the matrix of $T_1 + T_2$ is

$$\begin{aligned} A + B &= [a_{ij}] + [b_{ij}] \\ &= \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 3 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 4 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Example 5.35 : Let $A = [a_{ij}]$ be the matrix representation of a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ with respect to a basis $B = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$, with $\vec{u_1} = (5, 1, 3)$, $\vec{u_2} = (3, 2, 2)$ and $\vec{u_3} = (1, 2, 1)$, where

$$A = \begin{bmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 2 & 1 & -1 \end{bmatrix}$$

Find the linear transformation T.

Solution : For any $\vec{u} = (x_1, x_2, x_3) \in \mathbb{R}^3$ (domain), we have to find $T(\vec{u})$. From relation

$$T(\vec{u_i}) = a_{1i}\vec{u_1} + a_{2i}\vec{u_2} + a_{3i}\vec{u_3}, \quad i = 1, 2, 3$$

Since $\vec{u_i}$'s and a_{ij}'s are know we obtain

$T(\vec{u_1}) = T(5, 1, 3) = 3(5, 1, 3) + (-1)(3, 2, 2) + 2(1, 2, 1) = (14, 5, 9)$ … (i)

$T(\vec{u_2}) = T(3, 2, 2) = 2(5, 1, 3) + 0(3, 2, 2) + 1(1, 2, 1) = (11, 4, 7)$ … (ii)

$T(\vec{u_3}) = T(1, 2, 1) = (-2)(5, 1, 3) + 1(3, 2, 2) + (-1)(1, 2, 1) = (-8, -2, -5)$ … (iii)

Since B is basis for $\mathbb{R}^3$, any $\vec{u} = (x_1, x_2, x_3)$ can be written uniquely as

$$\begin{aligned}\vec{u} = (x_1, x_2, x_3) &= \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \alpha_3 \vec{u_3} \\ &= \alpha_1 (5, 1, 3) + \alpha_2 (3, 2, 2) + \alpha_3 (1, 2, 1) \\ &= (5\alpha_1 + 3\alpha_2 + \alpha_3, \alpha_1 + 2\alpha_2 + 2\alpha_3, 3\alpha_1 + 2\alpha_2 + \alpha_3)\end{aligned}$$

Equating the corresponding components, we get

$5\alpha_1 + 3\alpha_2 + \alpha_3 = x_1$ … (I)

$\alpha_1 + 2\alpha_2 + 2\alpha_3 = x_2$ … (II)

$3\alpha_1 + 2\alpha_2 + \alpha_3 = x_3$ … (III)

We solve these equations for α_1, α_2 and α_3.

Equation (I) and (III) gives

$2\alpha_1 + \alpha_2 = x_1 - x_3$ …(IV)

and equation (III) and (II) gives

$5\alpha_1 + 2\alpha_2 = 2x_3 - x_2$ … (V)

Next equation (V) and 2 (IV) gives

$\alpha_1 = -2x_1 - x_2 + 4x_3$

Using this value of α_1 in equation (IV), we obtain

$\alpha_2 = 5x_1 + 2x_2 - 9x_3$

Using equation (III) and values of α_1 and α_2, we get

$\alpha_3 = -4x_1 - x_2 + 7x_3$

Now, since T is linear, we have

$$T(\vec{u}) = \alpha_1 T(\vec{u_1}) + \alpha_2 T(\vec{u_2}) + \alpha_3 T(\vec{u_3})$$

Using equation (i), (ii) and (iii) and values of α_1, α_2, α_3 in terms of x_1, x_2 and x_3 we obtained, we get

$$\begin{aligned}T(\vec{u}) &= (-2x_1 - x_2 + 4x_3)(14, 5, 9) + (5x_1 + 2x_2 - 9x_3)(11, 4, 7) \\ &\quad + (-4x_1 - x_2 + 7x_3)(-8, -2, -5) \\ &= (59x_1 + 16x_2 - 99x_3, 18x_1 + 5x_2 - 30x_3, 37x_1 + 10x_2 - 62x_3)\end{aligned}$$

This is the formula for T.

CHANGE OF BASIS

We know that the co-ordinate vector of given vector in a vector space depends on the basis as well as the order of the basis. For example, if $\vec{u} = (2, 3)$ is a vector in $\mathbb{R}^2$, then the co-ordinate vector $\vec{x_1}$ of $\vec{u}$ with respect to the standard basis $\{(1, 0), (0, 1)\}$ is given by $\vec{x_1} = (2, 3)$. But the co-ordinate vector of $\vec{u}$ with respect to the basis $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}$ is $\left(\frac{5}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \vec{x_2}$. The question is what is relation between these two co-ordinates vectors $\vec{x_1}$ and $\vec{x_2}$. It is tempting thought that given a linear transformation T, when its matrix representation is simple and for this the change of basis is required.

Theorem 18 : $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be the two bases of a vector space of dimension n, then there exists a non-singular matrix $P = [p_{ij}]$ such that

$$\vec{v_i} = p_{1i}\vec{u_1} + p_{2i}\vec{u_2} + \ldots + p_{ni}\vec{u_n}, \quad i = 1, 2, \ldots, n$$

Proof : Since B is a basis for V, each basis vector $\vec{v_i}$ in B' can be expressed uniquely as a linear combination of vectors in B, that is

$$\vec{v_i} = p_{1i}\vec{u_1} + p_{2i}\vec{u_2} + \ldots + p_{ni}\vec{u_n}, \quad i = 1, 2, \ldots, n \qquad \ldots (i)$$

where, p_{ij} is are scalars. We define a matrix P, whose columns are $P_1, P_2, \ldots, P_n$, where,

$$P_i = \begin{bmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{bmatrix}$$

In fact it is to be noted that $P_1, P_2, \ldots, P_n$ are co-ordinate vectors of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}$ respectively with respect to the basis B, which are uniquely determined. This itself shows that P is non-singular. However, if for scalar $\alpha_1, \alpha_2, \ldots, \alpha_n$ and with the help of (i), we have

$$\alpha_1\vec{v_1} + \alpha_2\vec{v_2} + \ldots + \alpha_n\vec{v_n} = \alpha_1(b_{11}\vec{u_1} + \ldots + b_{n1}\vec{u_n}) + \alpha_n(b_{1n}\vec{u_1} + \ldots + b_{nn}\vec{u_n}) \ldots (ii)$$

Now, if $\alpha_1P_1 + \alpha_2P_2 + \ldots + \alpha_nP_n = \vec{0}$, then (ii) implies that

$\alpha_1p_{i1} + \ldots + \alpha_np_{in} = 0, \quad i = 1, 2, \ldots, n.$

Using this in equation (ii), we see that

$$\alpha_1\vec{v_1} + \ldots + \alpha_n\vec{v_n} = 0$$

Since B' is a basis this gives that $\alpha_1 = 0 = \ldots = \alpha_n$. So it follows that $P_1, P_2, \ldots, P_n$ are linearly independent, showing that P is non-singular matrix.

Remark : The relation (i) in the above theorem can be written in brief as

$$[\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n] = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n] P, \text{ the matrix form.}$$

Definition : The non-singular matrix P arising in the above theorem is called the **co-ordinate transformation matrix** or **transition matrix**.

Theorem 19 : Let $B = [\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ and $B' = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be two bases of a vector space V with dim V = n.

Let $P = [p_{ij}]$ be the transition matrix from B' to B, with

$$\vec{v}_i = p_{1i}\vec{u}_1 + p_{2i}\vec{u}_2 + \ldots + p_{ni}\vec{u}_n, \; i = 1, 2, \ldots, n \qquad \ldots (i)$$

$$X = [\alpha_1 \quad \alpha_2 \; \ldots \; \alpha_n]^t \text{ and } Y = [\beta_1 \quad \beta_2 \; \ldots \; \beta_n]^t$$

be the co-ordinate vectors of $\vec{u} \in V$ with respect to B and B' respectively, then

$$Y = P^{-1}X$$

Proof : Since X and Y are the co-ordinates of vectors of the same vector $\vec{u}$ with respect to B and B', we have

$$\begin{aligned}
\vec{u} &= \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots + \alpha_n\vec{u}_n \\
&= \beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \ldots + \beta_n\vec{v}_n \\
&= \sum_{i=1}^{n} \beta_i \left(\sum_{j=1}^{n} p_{ji}\vec{u}_j \right), && \text{by (i) in the statement} \\
&= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \beta_i p_{ji} \right) \vec{u}_j
\end{aligned}$$

Hence, we must have

$$\alpha_j = \sum_{i=1}^{n} \beta_i p_{ji}, \quad j = 1, 2, \ldots, n$$

This means $X = PY$

Since P is invertible $Y = P^{-1}X$.

Note :

(1) In the above proof X and Y are considered as column matrices.

(2) Since X and Y are co-ordinate vectors of the same vector, the above theorem shows that with help of the transition matrix B, given X we can find the vice versa, hence the name transition matrix.

Theorem 20 : Let $T : V \to V$ be a linear operator, when dim V = n.

Let $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be bases for V and let P be the transition matrix from B' and B. If $A = [a_{ij}]$ is the matrix representation of T with respect to the basis B, then the matrix representation of T with respect to B' is $P^{-1}AP$.

Before we prove then we remarks as :

Remark : If V is vector space and $B = \{\vec{u_1}, \ldots, \vec{u_n}\}$ is a basis for V. For a vector $\vec{u} \in V$, the co-ordinate vector of $\vec{u}$ with respect to the basis is denoted by $[\vec{u}]_B$.

Proof : Since P is the transition matrix from B' to B, for $\vec{u} \in V$, we have by above theorem

$$[\vec{u}]_B = P[\vec{u}]_{B'}$$

where, the j^{th} column of P is the co-ordinate vector $[\vec{v_j}]_{B'}$ of $\vec{v_j}$ with respect to B'.

Since P is invertible, this equation is also given as

$$[\vec{u}]_{B'} = P^{-1}[\vec{u}]_B \quad \ldots (I)$$

where, the j^{th} column of P^{-1} is the co-ordinate vector of $\vec{u_j}$ with respect to B.

Now, $A = [a_{ij}]$ is the matrix representation of T with respect to B, if we put $\vec{v} = T(\vec{u})$, then we have

$$[T(\vec{u})]_B = [\vec{v}]_B = A[\vec{u}]_B \quad \ldots (II)$$

Now, putting $\vec{v} = T(\vec{u})$ in equation (I), we get

$$[T(\vec{u})]_{B'} = P^{-1}[T(\vec{u})]_B$$

Now using first (II) and then the relation

$$[\vec{u}]_B = P[\vec{u}]_{B'}$$

In the last equation, we obtain

$$\begin{aligned} [T(\vec{u})]_{B'} &= P^{-1}A[\vec{u}]_B \\ &= P^{-1}A(P[\vec{u}]_{B'}) \quad \text{using (I)} \\ &= (P^{-1}AP)[\vec{u}]_{B'} \end{aligned}$$

Thus, $[T(\vec{u})]_{B'} = (P^{-1}AP)[\vec{u}]_{B'}$

This shows that $P^{-1}AP$ is the matrix representation of T with respect to B'.

Note : We visualize the above proof with the help of the following diagram.

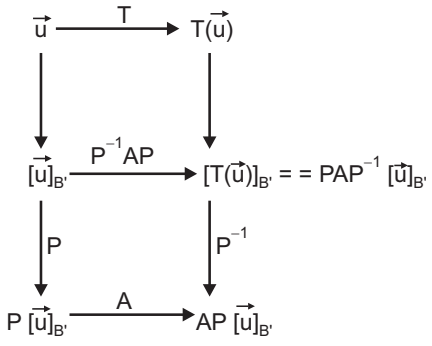


Fig. 5.2

Definition : Let A and B be two square matrices of order n, then **A is said to similar to B** if there exists a non-singular matrix P such that $B = P^{-1}AP$.

Note :

(1) If a square matrix A is said similar to matrix B, then

$$B = P^{-1}AP$$

for a non-singular matrix P. But this can also be written as

$$A = P^{-1}BP$$

Showing that B is similar to B : That is if A is similar to B then B is also similar to A, hence we may say that A and B are similar matrices.

(2) The determinants of two similar matrices are same. For, if A and B are similar matrices, then

$$B = P^{-1}AP$$

where, P is non-singular matrix.

Then
$$\begin{aligned} \det B &= \det (P^{-1}AP) \\ &= \det P^{-1} \cdot \det A \cdot \det P \\ &= \det A \left(\because \det P^{-1} = \frac{1}{\det P}, \det P \neq 0 \right) \end{aligned}$$

SOLVED EXAMPLES

Example 5.36 : Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation given by

$$T(x, y) = (x + y,\ x - 2y)$$

and let $B = \{(1, 0), (0, 1)$ and $B' = \{(1, -1), (2, 1)\}$

be bases for $\mathbb{R}^2$. Verify theorem (20).

Solution : First find $A = [a_{ij}]$, the matrix of T with respect to B, we have

$$\vec{u_1} = (1, 0),\ \vec{u_2} = (0, 1),\ \vec{v_1} = (1, -1),\ \vec{v_2} = (2, 1) \text{ and}$$

$$T(\vec{u_1}) = T(1, 0) = (1, 1) = 1 \cdot \vec{u_1} + 1 \cdot \vec{u_2}$$

$$T(\vec{u_2}) = T(0, 1) = (1, -2) = 1\vec{u_1} + (-2)\,\vec{u_2}$$

Thus, we have $A = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$

Now to find the transition matrix P, we have

$$\text{if } \vec{v_1} = \alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} \text{ and } \vec{v_2} = \beta_1 \vec{u_1} + \beta_2 \vec{u_2}$$

to find α_1, α_2 and β_1, β_2 the co-ordinate vectors of $\vec{v_1}$ and $\vec{v_2}$ with respect to B.

For this we have,

$$(1, -1) = \alpha_1 (1, 0) + \alpha_2 (0, 1) = (\alpha_1, \alpha_2)$$

$$(2, 1) = \beta_1 (1, 0) + \beta_2 (0, 1) = (\beta_1, \beta_2)$$

$$= (\alpha_1, \alpha_2)$$

which gives $\alpha_1 = 1,\ \alpha_2 = -1;\ \beta_1 = 2,\ \beta_2 = 1$, hence

$$P = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$

a transition matrix from B' to B.

Next the transition matrix P^{-1} from B to B' is given by

$$P^{-1} = \frac{1}{3}\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$$

Now, $$P^{-1}AP = \frac{1}{3}\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$

$$P^{-1}AP = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} \quad \ldots (I)$$

We calculate the matrix representation of T with respect to B' by direct calculations : We have

$$\left.\begin{aligned} T(\vec{v_1}) &= T(1, -1) = (0, 3) = a_{11}\vec{v_1} + a_{21}\vec{v_2} \\ \text{and} \quad T(\vec{v_2}) &= T(2, 1) = (3, 0) = a_{12}\vec{v_1} + a_{22}\vec{v_2} \end{aligned}\right\} \quad \ldots (II)$$

Using values of $\vec{v_1}$ and $\vec{v_2}$ and solving equations in (II), we obtain :

$$(0, 3) = (a_{11} + 2a_{21},\ -a_{11} + a_{21})$$

$$\Leftrightarrow \quad a_{11} + 2a_{21} = 0 \text{ and } -a_{11} + a_{21} = 3$$

This gives $a_{11} = -2a_{21}$, using second equation $2a_{21} + a_{21} = 3$ so $a_{21} = 1$ and $a_{11} = -2$.
And $(3, 0) = (a_{12} + 2a_{22}, -a_{12} + a_{22})$
This gives $a_{12} = 1$, $a_{22} = 1$.

Therefore, the matrix of T with respect to B' is $\begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$, which is same as $P^{-1}AP$ from equation (I), hence the theorem is verified.

Remark : In case of linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ (or $\mathbb{C}^n$ to $\mathbb{C}^m$) the matrix of T with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ (or $\mathbb{C}^n$ and $\mathbb{C}^m$) is called the **standard matrix of T.**

Exercise 5.2

1. Consider the basis $S = \{\vec{u_1}, \vec{u_2}\}$ for $\mathbb{R}^2$, where $\vec{u_1} = (1, 1)$ and $\vec{u_2} = (2, 0)$. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation for which $T(\vec{u_1}) = (-1, 3, 4)$, $T(\vec{u_2}) = (0, 2, -5)$. Find a formula for $T(x_1, x_2)$ and use it to compute $T(3, -6)$.
2. Let $S = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ be a basis for $\mathbb{R}^3$, where $\vec{u_1} = (1, 2, 3)$, $\vec{u_2} = (2, 5, 3)$ and $\vec{u_3} = (1, 0, 10)$. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation for which $T(\vec{u_1}) = (1, 0)$, $T(\vec{u_2}) = (1, 0)$, $T(\vec{u_3}) = (0, 1)$. Find the formula for $T(x, y, z)$ and use it to find $T(1, 1, 1)$.
3. Let $S = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ be basis for R^3, where $\vec{u_1} = (-1, 0, 1)$, $\vec{u_2} = (0, 1, -1)$ and $\vec{u_3} = (1, -1, 1)$. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation for which $T(\vec{u_1}) = \vec{e_1}$, $T(\vec{u_2}) = \vec{e_2}$ and $T(\vec{u_3}) = \vec{e_3}$, where $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\}$ is the standard basis for $\mathbb{R}^3$. Find a formula for $T(x, y, z)$ and use it compute $T(2, 1, -3)$.
4. Find the standard matrix for the linear transformation T in each case.
   (a) $T(x, y) = (2x - y, x + y)$
   (b) $T(x, y) = (x, y)$
   (c) $T(x, y, z) = (x + 2y + z, x + 5y, z)$
   (d) $T(x, y, z) = (4x_1, 7x_2, -8x_3)$
5. Find the standard matrix for the linear transformation T.
   (a) $T(x, y) = (y, -x, x + 3y, x - y)$
   (b) $T(x, y, z, w) = (7x + 2y - z + w, y + z - w)$
   (c) $T(x, y, z) = (0, 0, 0, 0, 0)$
   (d) $T(x, y, z, w) = (w, x, z, y, x - z)$
6. Find the standard matrix for T and use it to find the rank and nullity of T.
   (a) $T(x, y, z, w)$
   $= (x - 2y + z + 3w, 2x + y + 4w, -x - 3y + z - w, 3x + 3y + z + 7w)$
   (b) $T(x_1, x_2, x_3, x_4, x_5) = (x_1 + 4x_2 + 5x_3 + 9x_5,$
   $3x_1 - 2x_2 + x_3 - x_5, -x_1 - x_3 - x_5, 2x_1 + 3x_2 + 5x_3 + x_4 + 8x_5)$

7. If A and B are $m \times n$ matrices such that $AX = BX$ for all X in some basis for R^n, then prove that $A = B$.
8. Show that a map $\psi : M_{m \times n}(\mathbb{R}) \to \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ given by $\psi(A) = T_A$ is linear isomorphism.
9. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation given by
   $T(x, y, z) = (2x + 3y + z,\ x + 2y + 2z)$.
   Then show that the linear transformation
   $S : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $S(x, y) = (2x - 3y, -x + 2y, 0)$ is a right inverse of T.
10. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation defined by $T(x, y) = (2x + y, 3x + y)$. Show that :
    (a) T is bijective linear transformation, that is T is invertible linear transformation.
    (b) Find the inverse S of T.
11. Define a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(x, y) = \left(x - y, -\frac{1}{2}x, -\frac{1}{2}x + y\right)$.
    Find the left inverse S of T. Is the left inverse S of T is unique ? Justify.
12. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation. Let $B = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ and $B' = \{\vec{v_1}, \vec{v_2}\}$ be bases of $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively. If A is a matrix of T with respect to B and B', find the linear transform T, where, $A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & -2 \end{bmatrix}$.
13. Let A be the matrix representation of a linear transformation $T : P_1(x) \to P_2(x)$ with respect to the bases $B = \{1 - x, x\}$ and $B' = \{1, x, 1 + x + x^2\}$, where, $A = \begin{bmatrix} 2 & -1 \\ -1 & -1 \\ 1 & 1 \end{bmatrix}$. Find T.
14. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation given by
    $T(x_1, x_2, x_3) = (x_1 + 2x_2, 3x_3, x_2 - 2x_3)$.
    Find the matrix of T with respect to the bases :
    (a) $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} = B$ (b) $\{(1, 1, 1), (0, 1, 1), (0, 0, 1)\} = B'$.
15. In problem (14) above
    (a) Find the transition matrix P from B' to B.
    (b) Show that the matrices found in (a) and
    (b) In example (14) are similar.
16. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation given by;
    $T(x_1, x_2, x_3) = (x_1 + x_2 + x_3,\ 2x_2 + x_3)$.
    Let $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$, $B_1 = \{(2, 2, 1), (0, 1, 0), (1, 0, 1)\}$ be bases of $\mathbb{R}^3$ and
    $B' = \{(2, 1), (1, 1)\}$, $B_1' = \{(1, 0), (1, 1)\}$ be bases that of $\mathbb{R}^2$. Find (a) matrix A with respect to bases B and B' (b) matrix A_1 with respect to B_1 and B_1'.

Answers (5.2)

1. $T(x_1, x_2) = \left(-x_2, x_1 + 2x_2, -\frac{1}{2}x_1 + \frac{13}{2}x_2\right)$;

   $T(3, -6) = \left(6, -9, \frac{93}{2}\right)$

2. $T(x, y, z) = (30x - 10y - 3z, -9x + 3y + z)$, $T(1, 1) = (17, -5)$

3. $T(x, y, z) = (y + z, x + 2y + z, x + y + z)$

   $T(2, 1, -3) = (-2, 1, 0)$.

4. (a) $\begin{bmatrix} -2 & -1 \\ 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 1 & 2 & 1 \\ 1 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -8 \end{bmatrix}$

5. (a) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 1 & 3 \\ 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 7 & 2 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$

   (c) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix}$

6. (a) $\begin{bmatrix} 1 & -2 & 1 & 3 \\ 2 & 1 & 0 & 4 \\ -1 & -3 & 1 & -1 \\ 3 & 3 & 1 & 7 \end{bmatrix}$, rank (T) = 3, nullity (T) = 1

   (b) $\begin{bmatrix} 1 & 4 & 5 & 0 & 9 \\ 3 & -2 & 1 & 0 & -1 \\ -1 & 0 & -1 & 0 & -1 \\ 2 & 3 & 5 & 1 & 8 \end{bmatrix}$, rank (T) = 3, nullity = 2.

10. $S(u, v) = (v - u, 3u - 2v)$.

11. $S(u, v, w) = (2v, w + v)$ or $S(u, v, w) = (2v, 2v - u)$, S is not unique, since first answer show that for every u there is a left inverse of T, that is, there are infinitely many left inverse of T.

12. $T(x, y, z) = (5(9x + 3y - 16z), 3x - 4z)$.

13. For $p(x) = a_0 + a_1x$, $T(p(x) = 3a_0 + (2a_0 + a_1)x^2$.

14. (a) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 1 & -2 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 2 & 0 \\ 0 & 1 & 3 \\ -4 & -4 & -5 \end{bmatrix}$

15. $P = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

16. $A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \end{bmatrix}$, $A_1 = \begin{bmatrix} 0 & -1 & 1 \\ 5 & 2 & 1 \end{bmatrix}$

Important Points

- Definition, examples of linear transformation and its simple properties.
- Examples of transformations that are not linear.
- Codomain, domain, range or image of a linear transformation.
- Injective, surjective and bijective linear transformations, with examples.
- Sum, scalar multiple and composition of linear transformations.
- Kernel and range of linear transformation with illustrations.
- Dimension or rank-nullity theorem and its verification.
- Determination of linear transformation with prescribed values.
- Vector space of linear transformations from one vector space to another and its dimension theorem.
- Left and right inverse of a linear transformation, inverse of linear transformation.
- Necessary and sufficient conditions for existence of left, right and inverse of a linear transformation.
- Matrix of a linear transformation.
- Transition or co-ordinate transformation matrix.
- Given a matrix representation of a linear transformation T with respect to given bases to find T.
- Matrix of identity, zero, sum, scalar multiple, composition and inverse of a linear transformation.
- Change of basis and related theorems.

Miscellaneous Exericse

(A) Theory Questions :

1. Let $T : V \to W$ be a linear transformation, then prove that

   (a) $T(\vec{u} + \vec{u} + \ldots \text{ n-times}) = nT(\vec{u})$, n is positive integer.

(b) $T(-m\vec{u}) = mT(-\vec{u}) = -mT(\vec{u})$, m is positive integer.

(c) $T\left(\frac{p}{q}\vec{u}\right) = \frac{p}{q}T(\vec{u})$, where p and q are integers.

2. Let V and W be vector spaces over F and L(V, W) denote the set of all linear transformations from V to W, then prove that L(V, W) is a vector space over F under addition and scalar multiplication of linear transformation.

3. If dim V = m and dim W = n in problem (2), prove that dim (L(V, W)) = mn.

4. Let $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ be a basis of a vector space V and let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be arbitrary set of n-vectors in W. Then prove that there exists unique linear transformation $T : V \to W$ for which $T(\vec{u}_i) = \vec{v}_i$, $i = 1, 2, \ldots, n$.

5. Let $T : V \to W$ be a linear transformation. then prove that :

(a) ker T is a subspace of V.

(b) Im T = the range of T is subspace of W.

(c) ker T = $\{\vec{0}\}$ if and only if T is injective.

6. Prove that an injective linear transformation maps linearly independent sets into linearly independent sets, and vice versa.

7. State and prove dimension (or rank-nullity) theorem.

8. Prove that a linear transformation $T : V \to W$ is non-singular if and only if T maps basis of V into a basis of W.

9. Prove that a linear transformation $T : V \to W$ has an inverse if and only if it is bijective.

10. Let $T : V \to W$ be a linear transformation. Then T has

(a) a left inverse if and only if it is injective.

(b) a right inverse if and only if it is surjective.

11. Let $T : V \to W$ be a linear transformation between arbitrary vector spaces V and W, with $B = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ and $B' = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ as bases of V and W respectively. If $[\vec{u}]_B = (\alpha_1, \ldots, \alpha_n)$ is co-ordinate vector of $\vec{u}$ with respect to B and

$[\vec{v}]_{B'} = (\beta_1, \beta_2, \ldots, \beta_m)$ is the co-ordinate vector of $\vec{v}$ in W. If A is matrix representation of T with respect to the bases B and B'. Then prove that

$$T(\vec{u}) = \vec{v} \text{ if and only if } A[\vec{u}]_B = [\vec{v}]_{B'}.$$

12. Let T : V → W be linear transformation and let A be the matrix representation of T with respect to bases $B = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ of V and $B' = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}\}$ of W. Then prove that T is invertible if and only if A is invertible.

(B) Numerical Problems :

1. In each of the following maps, check whether the map is linear transformation. Justify you answer.
   (a) $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by $T(x, y, z) = (x, yz)$.
   (b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T(x, y, z) = (|x + y| + z, 2x - y, 3z)$.
   (c) $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(x, y) = (1, x + y)$.
   (d) $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T(x, y, z) = (x^2, x + y, x + z)$.
2. A linear transformation T is defined from $\mathbb{R}^2$ to $\mathbb{R}^3$ by $T(x, y) = (x - y, x + y)$. Show that T is bijective, that is, T is an isomorphism.
3. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation such that $T(\vec{u_1}) = (1, 0)$, $T(\vec{u_2}) = (0, 1)$, $T(\vec{u_3} = (1, 1)$, where $\vec{u_1} = (1, 1, -1)$, $\vec{u_2} = (4, 1, 1)$ and $\vec{u_3} = (1, -1, 2)$ be basis vectors of $\mathbb{R}^3$. Find T.
4. Find a linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ for which the null (kernel) space is spanned by $\vec{u_1} = (1, 1, 1, 1)$, $\vec{u_2} = (1, 0, 0, 1)$ and the image of T is spanned by $\vec{v_1} = (1, 1, 0)$ and $\vec{v_2} = (1, 0, 1)$.
5. Find a linear transformation in each of the following, which maps :
   (a) the vectors $\vec{u_1} = (2, 3)$, $\vec{u_3} = (3, 2)$ in $\mathbb{R}^2$ to $\vec{v_1} = (1, 2)$ and $\vec{v_2} = (2, 3)$ in $\mathbb{R}^2$ respectively.
   (b) the vectors $\vec{u_1} = (1, 2, 3)$, $\vec{u_2} = (2, 1, 3)$, $\vec{u_3} = (3, 1, 2)$ in $\mathbb{R}^3$ to $\vec{v_1} = (1, 0, 2)$, and $\vec{v_3} = (1, 2, 0)$ in $\mathbb{R}^3$ respectively.

6. Determine ker T and nullity of T, where T is given below :

   (a) $T(x, y, z) = (x - y, y - z, z - x)$.

   (b) $T(x, y, z) = x + y - z$.

   (c) $T(x, y, z) = (x + y, y + z)$.

7. Find the composite linear transformation of T_1 and T_2, $T_1 T_2$ and $T_2 T_1$, where,

   $T_1(x, y) = (y, x + y, x)$, $T_2(x, y, z) = (x - y, y - z)$

Answers

1. (a) Does not preserve addition and scalar multiplication.
   (b) Does not preserve addition and scalar multiplication.
   (c) Does not preserve addition.
   (d) Does not preserve addition and scalar multiplication.
3. $T(x, y, z) = (5x - 12y - 8z, x - 2y - z)$.
4. $T(x_1, x_2, x_3, x_4) = (x_3 + x_4 - x_1 - x_2, x_3 - x_2, x_4 - x_1)$.
5. (a) $T(x, y) = \frac{1}{5}(-y + 4x, 5x)$.

   (b) $T(x, y, z) = \frac{1}{6}(2x + 8y - 4z, 5x - y - z, -4x - 4y + 8z)$.
6. (a) ket T = straight line $x = y = z$ and nullity T = 1.
   (b) ker T = the plane orthogonal to the vector $(1, 1, -1)$ and passing through origin; nullity T = 2.
   (c) ker T = A straight line $x = y = z$, nullity T = 1.
7. $(T_1T_2)(x, y, z) = (y - z, x - z, x - y)$
   $(T_1T_2)(x, y) = (-x, y)$.

❐❐❐

Chapter **6**...

Inner Product Spaces

Synopsis

6.1 Introduction

6.2 Dot Product of Euclidean Inner Product in R^n

6.3 General Inner Product

6.4 Angle and Orthogonality in an Inner Product Space

6.5 Orthonormal Bases, Gram-Schmidt Process

6.6 Rank and Nullity : Sylvester Inequality

Objectives

(1) You will be able to understand the different types of product like dot product, inner product etc.

(2) You know the concept of orthonormal basis Rank and Nullity.

(3) You will be able to know the process, how to convert the given vectors into orthogonal and orhtonormal vector.

C. R. Rao

Calyampudi Radhakrishna Rao, FRS known as **C R Rao** (born 10 September 1920) is an Indian-born, naturalised American, mathematician and statistician. Rao has been honoured by numerous colloquia, honorary degrees, and festschrifts and was awarded the US National Medal of Science in 2002. The American Statistical Association has described him as "a living legend whose work has influenced not just statistics, but has had far reaching implications for fields as varied as economics, genetics, anthropology, geology, national planning, demography, biometry, and medicine." *The Times of India* listed Rao as one of the top 10 Indian scientists of all time. Rao is also a Senior Policy and Statistics advisor for the Indian Heart Association non-profit focused on raising South Asian cardiovascular disease awareness

He graduated, after an outstanding undergraduate career, with the Agrégation de Mathématiques in 1937. At this stage he decided that he would do his compulsory military service rather than delay it. He was assigned to the D.C.A. (Défense Contre Aéronefs - Defense Against Aircraft) but, being physically weak and lacking dexterity, he found that he was unable to perform tasks such as dismantling and assembling machine guns. He was posted first to Ballancourt, then to Biscarrosse, and although he had to extend his military service because of World War II, he never saw combat and was demobbed in August 1940. During this time, in 1938, he had married Marie-Hélène. In fact they had planned to marry in December 1935 but Marie-Hélène contracted a pulmonary tubercular infection and was sent to a sanatorium at Passy, Haut Savoie. The forced separation of eighteen months, during which time they could only correspond by letter, had been extremely difficult for both of them. After Schwartz left military service in 1940 he went with his wife to Toulouse where his parents had moved following the German invasion and the fall of France.

While in Toulouse, he met Henri Cartan when he visited there to conduct an oral on behalf of the École Normale Supérieure. In fact Marie-Hélène also took the opportunity to talk to Henri Cartan since she wanted to resume her mathematical studies. Cartan advised him to study for a doctorate at Clermont-Ferrand which is where the University of Strasbourg moved after the German armies invaded France at the start of World War II. Schwartz became a member of the Caisse National des Sciences (which later became the CNRS) which supported him until the end of 1942. After this support ended, he received funds from the ARS (Aid à la Recherche Scientifique), which supported him to the end of the war. Schwartz received mathematical advice from Georges Valiron who was based in Paris. He had known Valiron since attending his course *Functions of a complex variable* while he was an undergraduate.

6.1 Introduction

In this course, so far, we studied with exploring qualitative structures of a vector space and their elements without referring to any quantitative aspects of these spaces. The notion of length of a vector, and the distance and the angle between two vectors in a vector space is of great importance in this respect. We study these aspects in this chapter. We start with the known dot product of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, which we extend to $\mathbb{R}^n$, and thereby see the properties (theorem 1) satisfied by the dot product in $\mathbb{R}^n$. This theorem enables us to define axioms for the more general dot product, which we will call it as an inner product for vector space over $\mathbb{F}$. We should be aware that $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$ - the field of complex numbers. Although our geometric visualization does not extend beyond $\mathbb{R}^3$, interestingly it is possible to extend the familiar ideas like length of a vector, distance and angle between two vectors in a vector space over $\mathbb{F}$.

6.2 Dot Product of Euclidean Inner Product in $\mathbb{R}^n$

Let us recall the dot product in $\mathbb{R}^2$ or $\mathbb{R}^3$ and extend it to $\mathbb{R}^n$.

Definition : If $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ and θ is the angle between $\vec{u}$ and $\vec{v}$, then the dot product $\vec{u} \cdot \vec{v}$ is defined by;

$$\vec{u} \cdot \vec{v} = \begin{cases} |\vec{u}||\vec{v}| \cos\theta; & \text{if } \vec{u} \neq \vec{0} \text{ and } \vec{v} \neq \vec{0} \\ 0 & \text{; if either } \vec{u} = \vec{0} \text{ or } \vec{v} = \vec{0} \end{cases} \quad \dots (1)$$

Here, if $\vec{u} = (u_1, u_2, u_3)$ is a vector in $\mathbb{R}^3$, $|\vec{u}|$ denotes the length of $\vec{u}$ and is given by $|\vec{u}|^2 = u_1^2 + u_2^2 + u_3^2$.

Note : We use the notation $|\vec{u}|$ at the moment, however, we shall use the notation $\|\vec{u}\|$ after we define Euclidean inner product on $\mathbb{R}^n$.

Component Formula for the dot product :

Let $\vec{u} = (u_1, u_2, u_3)$ and $\vec{v} = (v_1, v_2, v_3)$ be two non-zero vectors. If, as shown in the Fig. 6.1, θ is the angle between $\vec{u}$ and $\vec{v}$, then the law of the cosines yields

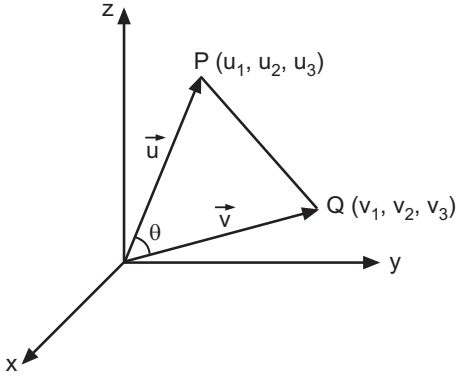


Fig. 6.1

$$|\overrightarrow{PQ}|^2 = |\vec{u}|^2 + |\vec{v}|^2 - 2|\vec{u}||\vec{v}|\cos\theta \qquad \ldots (2)$$

Since $\overrightarrow{PQ} = \vec{v} - \vec{u}$, we can rewrite (2) as

$$|\vec{u}||\vec{v}|\cos\theta = \frac{1}{2}(|\vec{u}|^2 + |\vec{v}|^2 - |\vec{v} - \vec{u}|^2)$$

or $$\vec{u} \cdot \vec{v} = \frac{1}{2}(|\vec{u}|^2 + |\vec{v}|^2 - |\vec{v} - \vec{u}|^2), \qquad \text{from (1).}$$

Substituting $|\vec{u}|^2 = u_1^2 + u_2^2 + u_3^2$, $|v|^2 = v_1^2 + v_2^2 + v_3^2$

and $|\vec{v} - \vec{u}|^2 = (v_1 - u_1)^2 + (v_2 - u_2)^2 + (v_3 - u_3)^2$, we get after simplifying

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + u_3v_3 \qquad \ldots (3)$$

If $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$ are two vectors in $\mathbb{R}^2$, then the corresponding formula is

$$\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 \qquad \ldots (4)$$

Euclidean n - space :

Definition : If $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ are any two vectors in $\mathbb{R}^n$, then the Euclidean inner product $\vec{u} \cdot \vec{v}$ is defined by

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

Observe that when n = 2 or n = 3, the Euclidean inner product is ordinary dot product formulae (3) and (4).

Example (1) : The Euclidean inner product of the vectors

$\vec{u} = (-1, 2, 3, 5)$ and $\vec{v} = (3, -2, 0, 4)$ in $\mathbb{R}^4$ is

$$\begin{aligned} \vec{u} \cdot \vec{v} &= (-1)(3) + (2)(-2) + (3)(0) + (5)(4) \\ &= -3 - 4 + 0 + 20 \\ &= 13. \end{aligned}$$

The vector space $\mathbb{R}^n$ with the usual operations of addition, scalar multiplication and the Euclidean inner product is called *Euclidean n-space.*

We prove the properties of Euclidean inner product in the following theorem.

Theorem 1 : If $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors in $\mathbb{R}^n$ and k is any scalar, then

(a) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$

(b) $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$

(c) $(k\vec{u}) \cdot \vec{v} = k(\vec{u} \cdot \vec{v})$

(d) $\vec{u} \cdot \vec{u} \geq 0$. Further $\vec{u} \cdot \vec{u} = 0$ if and only if $\vec{u} = 0$.

Proof : Let $\vec{u} = (u_1, u_2, \ldots, u_n)$, $\vec{v} = (v_1, v_2, \ldots, v_n)$ and $\vec{w} = (w_1, w_2, \ldots, w_n)$. Then

(a) $$\begin{aligned} \vec{u} \cdot \vec{v} &= u_1v_1 + u_2v_2 + \ldots + u_nv_n \\ &= v_1u_1 + v_2u_2 + \ldots + v_nu_n, \qquad \because\ u_iv_i = v_iu_i \\ &= \vec{v} \cdot \vec{u} \end{aligned}$$

(b) $$\begin{aligned} (\vec{u} + \vec{v}) \cdot \vec{w} &= (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \cdot (w_1, w_2, \ldots, w_n) \\ &= [(u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \ldots + (u_n + v_n)w_n] \\ &= u_1w_1 + v_1w_1 + u_2w_2 + v_2w_2 + \ldots + u_nw_n + v_nw_n \\ &= (u_1w_1, u_2w_2, \ldots, u_nw_n) + (v_1w_1, v_2w_2, \ldots, v_nw_n) \\ &= \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w} \end{aligned}$$

(c) $(k\vec{u}) \cdot \vec{v} = (ku_1, ku_2, \ldots, ku_n) \cdot (v_1, v_2, \ldots, v_n)$

$= (ku_1)v_1 + (ku_2)v_2 + \ldots + (ku_n)v_n$

$= k(u_1v_1 + u_2v_2 + \ldots + u_nv_n)$

$= k(\vec{u} \cdot \vec{v})$

(d) $\vec{u} \cdot \vec{u} = u_1^2 + u_2^2 + \ldots + u_n^2 \geq 0.$

Further, equality holds if and only if $u_1 = u_2 = \ldots = u_n = 0$, that is, $\vec{u} = 0$.

SOLVED EXAMPLES

Example 6.1 : Simplify $(2\vec{u} + \vec{v}) \cdot (3\vec{u} + 4\vec{v})$.

Solution : We have

$(2\vec{u} + \vec{v}) \cdot (3\vec{u} + 4\vec{v}) = (2\vec{u} + \vec{v}) \cdot (3\vec{u}) + (2\vec{u} + \vec{v}) \cdot (4\vec{v})$ by (b)

$= (2\vec{u}) \cdot (3\vec{u}) + \vec{v} \cdot (3\vec{u}) + (2\vec{u}) \cdot (4\vec{v}) + \vec{v} \cdot (4\vec{v})$ by (b) again

$= 6(\vec{u} \cdot \vec{u}) + 3(\vec{v} \cdot \vec{u}) + 8(\vec{u} \cdot \vec{v}) + 4(\vec{v} \cdot \vec{v})$ by (c)

$= 6(\vec{u} \cdot \vec{u}) + 11(\vec{u} \cdot \vec{v}) + 4(\vec{v} \cdot \vec{v})$ by (a)

Definition : *Euclidean norm* (or Euclidean length) of a vector $\vec{u} = (u_1, u_2, \ldots, u_n)$ in $\mathbb{R}^n$ is defined by

$$\| \vec{u} \| = (\vec{u} \cdot \vec{u})^{1/2} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

Similarly, the Euclidean distance between the vectors $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ is defined by

$$d(\vec{u}, \vec{v}) = \| u - v \| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$$

Alternative notions for vectors in $\mathbb{R}^n$

The vectors in $\mathbb{R}^n$ can be thought of as $n \times 1$ column matrices. For example, for $\vec{u}$ in $\mathbb{R}^n$, we can write $\vec{u}$ as

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_n \end{bmatrix}$$

The operations of addition and scalar multiplication can be worked with usual matrix addition and scalar multiplication of matrices.

With this notion for vectors in $\mathbb{R}^n$, if

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_n \end{bmatrix} \text{ and } \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ \vdots \\ v_n \end{bmatrix}, \text{ then it follows that } \vec{v}^t\vec{u} = [v_1, v_2, ..., v_n] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_n \end{bmatrix}$$

$= (u_1v_1 + u_2v_2 + ... + u_nv_n) = (\vec{u} \cdot \vec{v}) = \vec{u} \cdot \vec{v}$.

Thus, we have the matrix formula for the Euclidean inner product

$$\vec{v}^t \ \vec{u} = \vec{u} \cdot \vec{v}$$

SOLVED EXAMPLES

Example 6.2 : In each part, compute the Euclidean norm of the vectors :

(a) (1, –5), (b) (–2, 1, 2), (c) (1, 1, –2, 3, –4).

Solution :

(a) $\| (1, -5) \| = \sqrt{1^2 + (-5)^2} = \sqrt{26}$.

(b) $\| (-2, 1, 2) \| = \sqrt{(-2)^2 + 1^2 + (2)^2} = \sqrt{9} = 3.$

(c) $\| (1, 1, -2, 3, -4) \| = \sqrt{1^2 + 1^2 + (-2)^2 + 3^2 + (-4)^2} = \sqrt{31}$.

Example 6.3 : Let $\vec{u} = (1, 4, 3, 2)$, $\vec{v} = (3, 0, 8, -2)$ and $\vec{w} = (2, -2, 1, -3)$. Evaluate each of the followng expressions :

(a) $\| \vec{u} + \vec{v} \|$ (b) $\| \vec{u} \| + \| \vec{v} \|$ (c) $\| (-6) \vec{u} \| + 2 \| \vec{u} \|$

(d) $\| 2\vec{u} + (-3) \vec{v} + \vec{w} \|$ (e) $\dfrac{1}{\| \vec{u} \|} \vec{u}$ (f) $\left\| \dfrac{\vec{u}}{\| \vec{u} \|} \right\|$

Solution : (a) $\vec{u} + \vec{v} = (4, 4, 11, 0)$, so that

$$\| \vec{u} + \vec{v} \| = \| (4, 4, 11, 0) \| = \sqrt{4^2 + 4^2 + 11^2 + 0^2} = \sqrt{153} = 3\sqrt{17}$$

(b) $\| \vec{u} \| = \sqrt{1^2 + 4^2 + 3^2 + 2^2} = \sqrt{30}$

$\| \vec{v} \| = \sqrt{3^2 + 0^2 + 8^2 + (-2)^2} = \sqrt{77}$.

$\therefore \quad \| \vec{u} \| + \| \vec{v} \| = \sqrt{30} + \sqrt{77}$.

(c) $(-6)\,\vec{u} = (-6, -24, -18, -12)$

$$\therefore \quad \| (-6)\cdot \vec{u} \| = \sqrt{(-6)^2 + (-24)^2 + (-18)^2 + (-12)^2}$$
$$= \sqrt{1080} = 6\sqrt{30}$$

$$\therefore \quad \| (-6)\,\vec{u} \| + 2 \| \vec{u} \| = 6\sqrt{30} + 2\sqrt{30} = 8\sqrt{30}$$

Alternatively, $\| (-6)\,\vec{u} \| + 2 \| \vec{u} \|$

$$= 6 \| \vec{u} \| + 2 \| \vec{u} \|, \text{ by using (c) of theorem 1.}$$
$$= 8 \| \vec{u} \|$$
$$= 8\sqrt{30}\,.$$

(d) $\quad 2\vec{u} + (-3)\,\vec{v} + \vec{w} = (-5, 6, -17, 7)$

$$\therefore \quad \| 2\vec{u} + (-3)\,\vec{v} + \vec{w} \| = \sqrt{(-5)^2 + 6^2 + (-17)^2 + 7^2} = \sqrt{399}\,.$$

(e) $$\frac{1}{\| \vec{u} \|}\,\vec{u} = \frac{1}{\sqrt{30}}(1, 4, 3, 2) = \left(\frac{1}{\sqrt{30}}, \frac{4}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{2}{\sqrt{30}} \right)$$

(f) $$\left\| \frac{1}{\| \vec{u} \|}\,\vec{u} \right\| = \left\| \left(\frac{1}{\sqrt{30}}, \frac{4}{\sqrt{30}}, \frac{3}{\sqrt{30}}, \frac{2}{\sqrt{30}} \right) \right\|$$
$$= \sqrt{\frac{1}{30} + \frac{16}{30} + \frac{9}{30} + \frac{2}{30}} = 1.$$

6.3 General Inner Product

The length of a vector in $\mathbb{R}^n$ defined in previous section shows some anomalies in $\mathbb{C}^n$. In $\mathbb{C}$, if we have u = i, then by definition as in $\mathbb{R}^n$, the length of u is given by,

$$| u | = (i^2)^{1/2} = (-1)^{1/2} = i$$

which contradicts the notion of length which is always positive. To avoid this contradiction we define an inner product for vector space over $\mathbb{F}$.

Definition : Inner product or Scalar product :

Let V be arbitrary vector space over $\mathbb{F}$ and let $\vec{x}, \vec{y}, \vec{z}$ be any three vectors in V. The **inner product** of two vectors $\vec{x}$ and $\vec{y}$ in V denoted by $< \vec{x}, \vec{y} >$ is defined as a scalar in $\mathbb{F}$, which satisfies the following axioms :

1. $< \vec{x}, \vec{x} > \geq 0$ and $< \vec{x}, \vec{x} > = 0$ if and only if $\vec{x} = \vec{0}$. (Non-negativity)
2. $< \vec{x}, \vec{y} > = \overline{< \vec{y}, \vec{x} >}$. (Here the bar on the top denotes the complex conjugate). (Conjugate symmetry)
3. $< \alpha\vec{x} + \beta\vec{y}, \vec{z} > = \alpha < \vec{x}, \vec{z} > + \beta < \vec{y}, \vec{z} >$, where α and β are scalars in $\mathbb{F}$. (Linearity in first variable)

Definition : (Inner product space) :

The **inner product space** is a vector space in which the vectors satisfy Axioms 1 – 3, that is a vector space equipped with inner product is called an **inner product space.**

Note : $\mathbb{R}^n$ with an inner product is called the **Euclidean space** and $\mathbb{C}^n$ with an inner product is called **unitary space**. More generally, any **real vector space** V, for which an inner product is defined is called a Euclidean space. A complex vector space, for which an inner product is defined, is called a **unitary space**.

Remark :

(1) The inner product on a vector space V is a function $\langle , \rangle : V \times V \to \mathbb{F}$ satisfying the axioms 1 – 3.

(2) In case of real vector space the axiom (2) of conjugate symmetry becomes $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.

Example (1) : In $\mathbb{R}^n$, if $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ are two vectors, we define an inner product defined in earlier solution.

$$\langle \vec{u}, \vec{v} \rangle = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \sum_{i=1}^{n} u_iv_i$$

Then from theorem (1), and the fact that $\langle \vec{x}, \vec{y} \rangle$ is real number so its conjugate is itself, we see that this is an inner product on $\mathbb{R}^n$.

SOLVED EXAMPLES

Example 6.4 : Let $V = \mathbb{C}^n$, the complex vector space of n-tuples of complex numbers. For $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ define $\langle \vec{u}, \vec{v} \rangle$ as :

$$\langle \vec{u}, \vec{v} \rangle = u_1\bar{v_1} + u_2\bar{v_2} + \ldots + u_n\bar{v_n} = \sum_{i=1}^{n} u_i\bar{v_i} \quad \ldots (i)$$

where, $\bar{v_i}$ is complex conjugate of v_i. Then $\langle , \rangle$ defines an inner product on $\mathbb{C}^n$.

Solution : First observe that for any $\vec{u} = (u_1, u_2, \ldots, u_n)$ in $\mathbb{C}^n$,

$$\langle \vec{u}, \vec{u} \rangle = u_1\bar{u_1} + u_2\bar{u_2} + \ldots + u_n\bar{u_n}$$

$$= |u_1|^2 + |u_2|^2 + \ldots + |u_n|^2 \qquad (\because\ z\,\bar{z} = |z|^2)$$

Hence, $\langle \vec{u}, \vec{u} \rangle \geq 0, \quad \forall\ \vec{u} \in \mathbb{C}^n$

Next, we see that $\langle \vec{u}, \vec{u} \rangle = 0$ if and only if

$$|u_1|^2 = |u_2|^2 = \ldots = |u_n|^2 = 0$$

which implies that $u_1 = u_2 = \ldots = u_n = 0$, hence $\vec{u} = \vec{0}$. Thus axiom 1 is satisfied.

For second axiom, we see that

$$\langle \vec{u}, \vec{v} \rangle = \sum_{i=1}^{n} u_i \bar{v}_i = \sum_{i=1}^{n} \bar{u}_i v_i = \overline{\left(\sum_{i=1}^{n} \bar{u}_i v_i \right)} = \overline{\langle \vec{v}, \vec{u} \rangle} \quad \because \overline{(\bar{u}_i)} = u_i,\ i = 1, 2, \ldots, n$$

Next for any scalars α, β in $\mathbb{C}$, $\vec{u}, \vec{v}$ in $\mathbb{C}^n$, we have, for $\vec{w} = (w_1, w_2, \ldots, w_m) \in \mathbb{C}^n$, that

$$\begin{aligned} \langle \alpha \vec{u} + \beta \vec{v}, \vec{w} \rangle &= \sum_{i=1}^{n} (\alpha_i u_i + \beta_i v_i) \bar{w}_i, \text{ by definition (i)} \\ &= \sum_{i=1}^{n} (\alpha_i u_i \bar{w}_i + \beta_i v_i \bar{w}_i) = \sum_{i=1}^{n} \alpha (u_i \bar{w}_i) + \sum_{i=1}^{n} \beta (v_i \bar{w}_i) \\ &= \alpha \sum_{i=1}^{n} u_i \bar{w}_i + \beta \sum_{i=1}^{n} v_i \bar{w}_i = \alpha \langle \vec{u}, \vec{w} \rangle \beta \langle \vec{v}, \vec{w} \rangle \end{aligned}$$

All the Axiom 1-3 are satisfied, hence $\langle , \rangle$ is an inner product on $\mathbb{C}^n$.

Example 6.5 : For vectors $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$, and w_1, w_2, …, w_n any positive real numbers, define

$$\langle \vec{u}, \vec{v} \rangle = w_1 u_1 v_1 + w_2 u_2 v_2 + \ldots + w_n u_n v_n$$

Then this defines an inner product on $\mathbb{R}^n$ and is called **weighted inner product with weights $w_1, w_2, \ldots, w_n$.**

Solution : First for any $\vec{u} \in \mathbb{R}^n$.

$$\langle \vec{u}, \vec{u} \rangle = w_1 u_1^2 + w_2 u_2^2 + \ldots + w_n u_n^2, \text{ by definition.}$$

Since w_i's are positive it is clear that $\langle \vec{u}, \vec{u} \rangle \geq 0$.

Moreover $\langle \vec{u}, \vec{u} \rangle = 0$ if and only if $w_1 u_1^2 + w_2 u_2^2 + \ldots + w_n u_n^2 = 0$, this implies that $u_1 = 0 = u_2 = \ldots = u_n$.

$\therefore \vec{u} = \vec{0}$, thus axiom of non-negativity is satisfied.

Next we are working with real vector space, and by definition it is obvious that

$$\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$$

and conjugate symmetry follows.

Now, $\vec{u} = (u_1, u_2, \ldots, u_n)$, $\vec{v} = (v_1, v_2, \ldots, v_n)$, $\vec{w} = (w_1, w_2, \ldots, w_n)$ in $\mathbb{R}^n$ and for any scalars (real) α, β, we have;

$$\begin{aligned} \langle \alpha \vec{u} + \beta \vec{v}, \vec{w} \rangle &= \sum_{i=1}^{n} (\alpha \cdot u_i + \beta \cdot v_i) w_i), \text{ by definition} \\ &= \sum_{i=1}^{n} \alpha u_i w_i + \sum_{i=1}^{n} \beta v_i w_i \\ &= \alpha \sum_{i=1}^{n} u_i w_i + \beta \sum_{i=1}^{n} v_i w_i = \alpha \langle \vec{u}, \vec{w} \rangle + \beta \langle \vec{v}, \vec{w} \rangle \end{aligned}$$

Example 6.6 : Let $\vec{u} = (u_1, u_2, u_3)$ and $\vec{v} = (v_1, v_2, v_3)$ be vectors in $\mathbb{R}^3$. Define $< \vec{u}, \vec{v} > = 3\, u_1 v_1 + 2\, u_2 v_2 + u_1 v_1$.

Show that, this defines an inner product with weights 3, 2 and 1 on $\mathbb{R}^3$.

Solution : This is a particular case of example (2). (The students are requested to verify the four axioms as an exercise).

Example 6.7 : For $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2)$ in $\mathbb{R}^2$,

define $< \vec{x}, \vec{y} > = y_1 (x_1 + 2x_2) + y_2 (2x_1 + 5x_2)$. … (i)

Show that $< , >$ defines an inner product.

Solution : (1) Symmetry : From definition (i), we have,

$$< \vec{x}, \vec{y} > = y_1 x_1 + 2x_2 y_1 + 2y_2 x_1 + 5y_2 x_2 \quad \text{… (ii)}$$

$\therefore$ $$< \vec{y}, \vec{x} > = x_1 (y_1 + 2y_2) + x_2 (2y_1 + 5y_2)$$
$$= x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 5x_2 y_2 \quad \text{… (iii)}$$

From (ii) and (iii) $< \vec{x}, \vec{y} > = < \vec{y}, \vec{x} >$

(2) Let $\vec{z} = (z_1, z_2)$ be any vector in $\mathbb{R}^2$. Then

$$\vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2) \text{ by definition of addition in } \mathbb{R}^2.$$

By definition (i), we have

$$< \vec{x} + \vec{y}, \vec{z} > = z_1 [(x_1 + y_1) + 2 (x_2 + y_2)] + z_2 [2 (x_1 + y_1) + 5 (x_2 + y_2)]$$
$$= z_1 x_1 + z_1 y_1 + 2z_1 x_2 + 2z_1 y_2 + 2z_2 x_1 + 2z_2 y_1$$
$$+ 5z_2 x_2 + 5z_2 y_2 \quad \text{… (iv)}$$

Also, we compute from definition (i) that

$$< \vec{x}, \vec{y} > + < \vec{y}, \vec{z} > = z_1 (x_1 + 2x_2) + z_2 (2x_1 + 5x_2) + z_1 (y_1 + 2y_2) + z_2 (2y_1 + 5y_2)$$
$$= z_1 x_1 + 2z_1 x_2 + 2z_2 x_1 + 5z_2 x_2$$
$$+ z_1 y_1 + 2y_1 y_2 + 2z_2 y_1 + 5z_2 y_2 \quad \text{… (v)}$$

Comparing r.h.s. of (iv) and (v) we see that

$$< \vec{x} + \vec{y}, \vec{z} > = < \vec{x}, \vec{z} > + < \vec{y}, \vec{z} >$$

(3) For any scalar α, $\alpha \vec{x} = (\alpha x_1, \alpha x_2)$, so that

$$< \alpha \vec{x}, \vec{y} > = y_1 (\alpha x_1 + 2\alpha x_2) + y_2 (2\alpha x_1 + 5\alpha x_2)$$
$$= \alpha [y_1 (x_1 + 2x_2) + y_2 (2x_1 + 5x_2)]$$
$$= \alpha < \vec{x}, \vec{y} >$$

which establishes homogeneity of $< , >$.

(4) **Non-negativity :** We have from definition

$$\begin{aligned} < \vec{x} , \vec{x} > &= x_1 (x_1 + 2x_2) + x_2 \ (2x_1 + 5x_2) \\ &= x_1^2 + 2x_1 x_2 + 2x_1 x_2 + 5x_2^2 \\ &= x_1^2 + 4x_1 x_2 + 5x_2^2 \\ &= (x_1 + 2x_2)^2 + x_2^2 \ \geq 0, \text{ for } \vec{x} \in \mathbb{R}^2 \end{aligned}$$

Also, $< \vec{x} , \vec{x} > = 0$ if and only if

$(x_1 + 2x_2)^2 + x_2^2 = 0$

This means $(x_1 + 2x_2) = 0$ and $x_2 = 0$ hence $x_1 = 0$ and $x_2 = 0$

Thus, $< \vec{x} , \vec{x} > = 0$ if and only if $\vec{x} = \vec{0}$.

Thus, $< , >$ defines an inner product on $\mathbb{R}^2$.

Example 6.8 : Let

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ : \\ : \\ u_n \end{bmatrix} \text{ and } \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ : \\ : \\ v_n \end{bmatrix}$$

be the vectors in $\mathbb{R}^n$ (expressed as $n \times 1$ matrices), and let A be an invertible $n \times n$ matrix. If $\vec{u} \cdot \vec{v}$ is the Euclidean inner product on $\mathbb{R}^n$, then the formula

$$< \vec{u}, \vec{v} > = A\vec{u} \cdot A\vec{v} \qquad \ldots (*)$$

is an inner product on $\mathbb{R}^n$, it is called the *inner product on $\mathbb{R}^n$ generated by A.*

Solution : (1) Since $A\vec{u}$ and $A\vec{v}$ are column vectors and $A\vec{u} \cdot A\vec{v}$ is the Euclidean inner product, it is clear that

$$\begin{aligned} < \vec{u}, \vec{v} > &= A\vec{u} \cdot A\vec{v} \\ &= A\vec{v} \cdot A\vec{u} = < \vec{v}, \vec{u} > \end{aligned}$$

(2) For any $\vec{u}$, $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$,

$$\begin{aligned} < \vec{u} + \vec{v}, \vec{w} > &= A(\vec{u} + \vec{v}) \cdot A\vec{w} \\ &= (A\vec{u} + A\vec{v}) \cdot A\vec{w} \\ &= (A\vec{u} \cdot A\vec{w}) + (A\vec{v} \cdot A\vec{w}) \\ &= < \vec{u}, \vec{w} > + < \vec{v}, \vec{w} > \end{aligned}$$

(3) For any scalar k,

$$\begin{aligned} \langle k\vec{u}, \vec{v} \rangle &= A(k\vec{u}) \cdot A\vec{v} \\ &= (k(A\vec{u})) \cdot A\vec{v} \\ &= k(A\vec{u} \cdot A\vec{v}) \\ &= k\langle \vec{u}, \vec{v} \rangle \end{aligned}$$

(4) $\langle \vec{u}, \vec{v} \rangle = A\vec{u} \cdot A\vec{v} \geq 0$, by (d) of theorem 1. Further, $\langle \vec{u}, \vec{u} \rangle = 0$ iff $A\vec{u} = \vec{0}$. And $A\vec{u} = \vec{0}$ iff $\vec{u} = \vec{0}$, since A is invertible matrix and $A\vec{u} = \vec{0}$ is homogeneous system and has only trivial solution.

Remark : As we know that the Euclidean inner product $\vec{u} \cdot \vec{v}$ can be written as the matrix formula $\vec{v}^t \vec{u}$, it follows that (*) can be written in the alternative form as,

$$\langle \vec{u}, \vec{v} \rangle = (A\vec{v})^t (A\vec{u})$$

or equivalently, $$\langle \vec{u}, \vec{v} \rangle = \vec{v}^t A^t A\vec{u} \quad \text{... (**)}$$

Example 6.9 : The inner product on $\mathbb{R}^n$ generated by the $n \times n$ identity matrix is the Euclidean inner product, since substituting $A = I$ in (*) yields

$$\langle \vec{u}, \vec{v} \rangle = I\vec{u} \cdot I\vec{v} = \vec{u} \cdot \vec{v}$$

The weighted inner product $\langle \vec{u}, \vec{v} \rangle = 3u_1v_1 + 2u_2v_2 + u_3v_3$ discussed in example (3) is the inner product on $\mathbb{R}^3$ generated by

$$A = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

because substituting this matrix in (**) yields

$$\begin{aligned} \langle \vec{u}, \vec{v} \rangle &= [v_1, v_2, v_3] \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \\ &= [v_1, v_2, v_3] \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \\ &= 3u_1v_1 + 2u_2v_2 + u_3v_3. \end{aligned}$$

In general, the weighted Euclidean inner product

$$\langle \vec{u}, \vec{v} \rangle = w_1u_1v_1 + w_2u_2v_2 + \ldots\ldots + w_nu_nv_n$$

is the inner product on $\mathbb{R}^n$ generated by

$$A = \begin{bmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ : & : & : & & : \\ : & : & : & & : \\ 0 & 0 & 0 & & \sqrt{w_n} \end{bmatrix}$$

Remark : Euclidean weighted inner product on $\mathbb{R}^n$ is the particular case of the inner product generated by A on $\mathbb{R}^n$.

Example 6.10 : Let f = f (x) and g = g (x) be two functions in the vector space V of continuous functions defined on [0, 1] (one can take a and b in place of 0 and 1 respectively) and define

$$< f, g > = \int_0^1 f(x)\, g(x)\, dx.$$

Show that this formula defines an inner product on V.

Solution :

(1) $< f, g > = \int_0^1 f(x)\, g(x)\, dx = \int_0^1 g(x)\, f(x)\, dx = < g, f >$

(2) For f, g, and h in V,

$$< f + g, h > = \int_0^1 (f(x) + g(x))\ h(x)\, dx = \int_0^1 (f(x)\, h(x) + g(x)\, h(x))\ dx$$

$$= \int_0^1 f(x)\, h(x)\, dx + \int_0^1 g(x)\, h(x)\, dx$$

by properties of definite integrals

$$= < f, h > + < g, h >$$

(3) For any scalar k, we have

$$< kf, g > = \int_0^1 (k\, f(x))\ g(x)\, dx$$

$$= \int_0^1 k\, (f(x)\, g(x))\, dx = k \int_0^1 f(x)\, g(x)\, dx$$

$$= k < f, g >.$$

(4) $$< f, f > = \int_0^1 f(x)\, f(x)\, dx = \int_0^1 [f(x)]^2\, dx \geq 0, \text{ since } [f(x)]^2 \geq 0$$

Further equality holds iff f (x) = 0 for all x | [0, 1], because f is continuous in [0, 1], that is f = 0.

Example 6.11 : Let V = C [a, b] be the vector space of all continuous real valued functions defined on [a, b]. For f, g in V, define

$$< f, g > = \int_a^b f(x)\, g(x)\, dx$$

Then it can be seen that this is an inner product on [a, b], which is generalization of the example (9). Also note that this is **real vector space**.

Example 6.12 : Let V = $P_n(x)$, the vector space of all polynomials of degree ≤ n over $\mathbb{C}$. For any two polynomials p(x) and q(x) in $P_n(x)$, define

$$< p, q > = \int_a^b p(x)\, \overline{q(x)}\, dx$$

where, a and b are any scalars in $\mathbb{C}$ with $a \neq b$, then this defines an inner product.

Solution : First we note that every polynomial p(x) can be written as $p(x) = p_1(x) + ip_2(x)$, where $P_1(x)$, $P_2(x)$ are polynomials over $\mathbb{R}$ and $i = \sqrt{-1}$, similarly, $q(x) = q_1(x) + iq_2(x)$.

Then we have,

$$< p(x), p(x) > = \int_a^b p(x) \cdot \overline{p(x)}\, dx = \int_a^b (p_1(x) + ip_2(x))\, \overline{(p_1(x) + ip_2(x))}\, dx$$

$$= \int_a^b [\{p_1(x)\}^2 + \{p_2(x)\}^2]\, dx$$

So we see that $< p(x), p(x) > \geq 0$ and $< p(x), p(x) > = 0$ if and only if $p_1(x) = 0$ and $p_2(x) = 0$; that $p(x) = 0$. So the first axiom is satisfied.

Next for any p, q in V, we have

$$< p, q > = \int_a^b p(x)\, \overline{q(x)}\, dx = \int_a^b \overline{\overline{p(x)}\, q(x)}\, dx$$

$$= \left(\int_a^b q(x)\, \overline{p(x)}\, dx \right)$$

$$= < \overline{q(x), p(x)} >$$

Thus, the conjugate symmetry follows.

Again for any scalars α, β in $\mathbb{C}$ and for polynomials in p, q, r in V, we have

$$< \alpha p + \beta q, r > = \int_a^b (\alpha p + \beta q)(x) \cdot \overline{r(x)}\, dx$$

$$= \int_a^b [\alpha\, p(x)\, \overline{r(x)} + \beta\, q(x)\, \overline{r(x)}]\, dx$$

$$= \alpha \int_a^b p(x)\, \overline{r(x)}\, dx + \beta \int_a^b q(x)\, \overline{r(x)}\, dx$$

$$= \alpha < p(x), r(x) > + \beta < q(x), r(x) >$$

Example 6.13 : Let V be the vector space of all continuous functions defined on the internal $[0, 2\pi]$ with inner product defined by

$$< f, g > = \int_0^{2\pi} f(x) \cdot \overline{g(x)}\, dx$$

where, f and g are two elements of V, then V is an inner product space.

Solution : For any $f \in V$, it can be written as $f = f_1 + i f_2$, $i = \sqrt{-1}$ and f_1, f_2 are real valued continuous functions defined on $[0, 2\pi]$.

Therefore,
$$< f, f > = \int_0^{2\pi} f(x) \cdot \overline{f(x)}\, dx$$
$$= \int_0^{2\pi} [f_1(x) + i f_2(x)] \ \overline{[f_1(x) + i f_2(x)]}\, dx$$
$$= \int_0^{2\pi} [\{f_1(x)\}^2 + \{f_2(x)\}^2]\, dx$$

and we see that $< f, f > > 0$ if and only if $f \neq 0$ in some subinterval of $[0, 2\pi]$ and $< f, f > = 0$ if and only if $f = 0$.

The conjugate symmetry and linearity in first component can be verified similar to that in example (11), thus $<, >$ defined is an inner product on V, therefore V is an inner product space.

Now, we prove the following basic algebraic properties of inner products.

Theorem (2) : Properties of inner product :

Let V be an inner product space. For $\vec{u}$, $\vec{v}$ and $\vec{w}$ vectors in V and α, β and scalars then

(a) $< \vec{u}, \vec{0} > = < \vec{0}, \vec{u} > = 0$.

(b) $< \vec{u}, \alpha \vec{v} + \beta \vec{w} > = \bar{\alpha} < u, \vec{v} > + \bar{\beta} < \vec{u}, \vec{w} >$.

Proof : (a) For $\vec{0} \in V$, we have $\vec{0} + \vec{0} = \vec{0}$, hence

$$< \vec{0}, \vec{u} > = < \vec{0} + \vec{0}, \vec{u} >$$
$$= < \vec{0}, \vec{u} > + < \vec{0}, \vec{u} >, \text{ additivity in first component.}$$

Adding $- < \vec{0}, \vec{u} >$ on both sides we obtain $< \vec{0}, \vec{u} > = 0$.

Also $< \vec{u}, \vec{0} > = \overline{< \vec{0}, \vec{u} >}$, conjugate symmetry by above

$$= \bar{0} = 0$$

$\therefore$ $< \vec{u}, \vec{0} > = 0$

(b) We have,

$$\langle \vec{u}, \alpha\vec{v} + \beta\vec{w} \rangle = \overline{\langle \alpha\vec{v} + \beta\vec{w}, \vec{u} \rangle}, \text{ conjugate symmetry.}$$

$$= \overline{[\alpha \langle \vec{v}, \vec{u} \rangle + \beta \langle \vec{w}, \vec{u} \rangle]} \qquad \text{by axiom 3}$$

$$= \bar{\alpha} \overline{\langle \vec{v}, \vec{u} \rangle} + \bar{\beta} \overline{\langle \vec{w}, \vec{u} \rangle}$$

$$= \bar{\alpha} \langle \vec{u}, \vec{v} \rangle + \bar{\beta} \langle \vec{u}, \vec{w} \rangle$$

Note that in case of real vector space α, β being real numbers (b) takes the form :

$$\langle \vec{u}, \alpha\vec{v} + \beta\vec{w} \rangle = \alpha \langle \vec{u}, \vec{v} \rangle + \beta \langle \vec{u}, \vec{w} \rangle$$

Norm or length of a vector in an inner product space :

Definition : Let V be an inner product space, and $\vec{u} \in V$, then **norm (or length) of $\vec{u}$,** denoted by $\| \vec{u} \|$ is defined as

$$\| \vec{u} \| = \{ \langle \vec{u}, \vec{u} \rangle \}^{1/2} = \sqrt{\langle \vec{u}, \vec{u} \rangle}$$

For two vectors $\vec{u}$ and $\vec{v}$ in V, we define the distance between $\vec{u}$ and $\vec{v}$, denoted by $d(\vec{u}, \vec{v})$ as

$$d(\vec{u}, \vec{v}) = \| \vec{u} - \vec{v} \|$$

Note that the norm and hence also distance depends on the inner product defined on a vector space.

SOLVED EXAMPLES

Example 6.14 : Find norm of each and distance between

$\vec{u} = (1, 1, -1)$ and $\vec{v} = (-1, 1, 0)$ in $\mathbb{R}^3$ with respect to the

(a) usual Euclidean dot product.

(b) inner product defined by :

For $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$, define

$$\langle \vec{u}, \vec{v} \rangle = u_1v_1 + 2u_2v_2 + 3u_3v_3 \qquad \text{(weighted I.P.)}$$

Solution : (a) In case of Euclidean inner product on $\mathbb{R}^3$, we have,

$$\| \vec{u} \| = \sqrt{\langle \vec{u}, \vec{u} \rangle} = \sqrt{1 \cdot 1 + 1 \cdot 1 + (-1)(-1)} = \sqrt{3}$$

and $$\| \vec{v} \| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{(-1)(-1) + 1 \cdot 1 + 0 \cdot 0} = \sqrt{2}$$

As, $$\vec{u} - \vec{v} = (1, 1, -1) - (-1, 1, 0) = (2, 0, -1)$$

$\therefore$ The distance between $\vec{u}$ and $\vec{v}$ is

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \sqrt{(2\cdot 2) + 0\cdot 0 + (-1)(-1)} = \sqrt{5}$$

(b) In this case we have for $\vec{u} = (u_1, u_2, u_3)$

$$\|\vec{u}\| = \sqrt{u_1^2 + 2u_2^2 + 3u_3^2}$$

Hence, $$\|\vec{u}\| = \|(1, 1, -1)\| = \sqrt{1\cdot 1 + 2\cdot(1\cdot 1) + 3(-1)(-1)} = \sqrt{6}$$

and $$\|\vec{v}\| = \|(-1, 1, 0)\| = \sqrt{(-1)(-1) + 2\cdot(1\cdot 1) + 3(0\cdot 0)} = \sqrt{3}$$

and the distance between the vector $\vec{u}$ and $\vec{v}$

$$d(\vec{u}, \vec{v}) = \|(2, 0, -1)\| = \sqrt{2\cdot 2 + 2(0\cdot 0) + 3(-1)(-1)} = \sqrt{7}$$

Example 6.15 : Find $<\vec{u}, \vec{v}>, <\vec{v}, \vec{u}>$ and norm of each $\vec{u} = (2, 2i, -1)$, $\vec{v} = (2, -1, 1+i)$ in $\mathbb{C}^3$ with respect to the inner product defined in example (2) with n = 3. Also find $d(\vec{u}, \vec{v})$.

Solution : We have, for $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$

$$<\vec{u}, \vec{v}> = u_1\bar{v}_1 + u_2\bar{v}_2 + u_3\bar{v}_3$$

Therefore with $\vec{u} = (2, 2i, -1)$, $\vec{v} = (2, -1, 1+i)$

$$\begin{aligned} <\vec{u}, \vec{v}> &= 2\cdot\bar{2} + 2i\,(\overline{-1}) + (-1)\,(\overline{1+i}) \\ &= 2\cdot 2 + (2i)(-1) + (-1)(1-i) \\ &= 4 - 2i - 1 + i \\ &= 3 - i \end{aligned}$$

and $$\begin{aligned} <\vec{v}, \vec{u}> &= 2\cdot\bar{2} + (-1)\,(\overline{2i}) + (1+i)\,(\overline{-1}) \\ &= 2\cdot 2 + (-1)(-2i) + (1+i)(-1) \\ &= 4 + 2i - 1 - i = 3 + i \end{aligned}$$

Next, $$\begin{aligned} \|\vec{u}\| &= \sqrt{<\vec{u}, \vec{v}>} \\ &= \sqrt{2\cdot\bar{r} + (2i)\,(\overline{2i}) + (-1)\,(\overline{-1})} \\ &= \sqrt{2\cdot 2 + (2i)(-2i) + (-1)(-1)} \\ &= \sqrt{4 + 4 + 1} \\ &= \sqrt{9} \end{aligned}$$

and $$\|\vec{v}\| = \sqrt{2 \cdot \bar{2} + (-1)(\overline{-1}) + (1+i)(\overline{1+i})}$$
$$= \sqrt{2 \cdot 2 + (-1)(-1) + (1+i)(1-i)} \qquad \because \overline{1+i} = 1-i$$
$$= \sqrt{4+1+2}$$
$$= \sqrt{7}$$

Now as $$\vec{u} - \vec{v} = (0, 2i+1, -2-i)$$
$$d(\vec{u}, \vec{v}) = \sqrt{0 \cdot \bar{0} + (2i+1)(\overline{2i+1}) + (-2-i)(\overline{-2-i})}$$
$$= \sqrt{0+5+5}$$
$$= \sqrt{10}$$

Definition : Let V be an inner product space and $\vec{u} \in V$, then $\vec{u}$ is called a **unit vector** in V if $\|\vec{u}\| = 1$. The set of vectors $\vec{u}$ in V satisfying $\|\vec{u}\| = 1$ is called the **unit sphere** or the **unit circle** in V.

Remark : Given a vector $\vec{u}$ in an inner product space V, we can **normalize** $\vec{u}$ if it is divided by its norm, then after normalizing $\vec{u}$, we have $\dfrac{\vec{u}}{\|\vec{u}\|}$ which is of norm 1. Thus, by normalizing a vector we can obtain a unit vector.

For example, $\vec{u} = (-1, 1, 2)$ in $\mathbb{R}^3$, with usual Euclidean inner product, $\|\vec{u}\| = \sqrt{6}$ and $\vec{x} = \dfrac{\vec{u}}{\|\vec{u}\|} = \dfrac{1}{\sqrt{6}}(-1, 1, 2)$ is a normalized vector obtain form $\vec{u}$ and clearly $\|\vec{x}\| = 1$.

Remark :

(1) In an inner product space V, the norm of a vector satisfies the following :

(a) $\|\vec{u}\| \geq 0$, for all $\vec{u} \in V$ and $\|\vec{u}\| = 0$ iff $\vec{u} = \vec{0}$.

(b) $\|\alpha \vec{u}\| = |\alpha| \, \|\vec{u}\|$, α a scalar (homogeneity).

(c) $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$, for $\vec{u}, \vec{v} \in V$, this is called triangle inequality.

(2) Similarly distance function defined on $V \times V$, satisfied.

(a) $d(\vec{u}, \vec{v}) \geq 0$, for $\vec{u}, \vec{v} \in V$ and $d(\vec{u}, \vec{v}) = 0$ iff $\vec{u} = \vec{v}$.

(b) $d(\vec{u}, \vec{v}) = d(\vec{v}, \vec{u})$ (symmetry).

(c) $d(\vec{u}, \vec{v}) \leq d(\vec{u}, \vec{w}) + d(\vec{w}, \vec{v})$ (triangle inequality) for any $\vec{w} \in V$.

All these results (axioms) can be proved by using the axioms of inner product.

SOLVED EXAMPLES

Example 6.16 : (a) Sketch the unit circle in an XY - plane using Euclidean inner product $< \vec{u}, \vec{v} > = u_1v_1 + u_2v_2$.

(b) Sketch the unit circle in an XY-plane using the weighted Euclidean inner product $< \vec{u}, \vec{v} > = \frac{1}{16} u_1v_1 + \frac{1}{9} u_2v_2$.

Solution : (a) If $\vec{u} = (x, y)$, then $\| \vec{u} \| = < \vec{u}, \vec{u} >^{1/2} = \sqrt{x^2 + y^2}$, on squaring both sides,

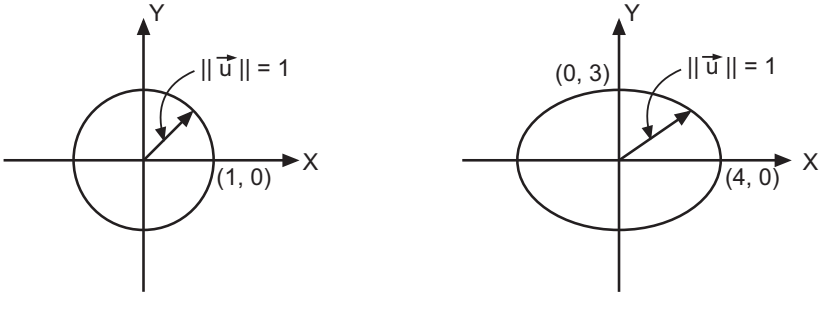
$$x^2 + y^2 = 1 \text{ which is a unit standard circle.}$$

So, the graph of this equation is a circle of radius 1 centred at the origin.

(b) If $\vec{u} = (x, y)$, then $\| \vec{u} \| = < u, u >^{1/2} = \sqrt{\frac{1}{16} x^2 + \frac{1}{9} y^2}$, so the equation is $\sqrt{\frac{1}{16} x^2 + \frac{1}{9} y^2} = 1$, or on squaring both sides,

$$\frac{x^2}{16} + \frac{y^2}{9} = 1.$$

The graph of this equation is the ellipse.



(a) The Unit Circle with the Euclidean Norm

$$\| \vec{u} \| = \sqrt{x^2 + y^2}$$

(b) The Unit Circle with the Norm

$$\| \vec{u} \| = \sqrt{\frac{1}{16} x^2 + \frac{1}{9} y^2}$$

Fig. 6.2

Example 6.17 : Suppose that $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors in a real vector space V such that $< \vec{u}, \vec{v} > = -1, < \vec{v}, \vec{w} > = 3, < \vec{u}, \vec{w} > = 2, \| \vec{u} \| = 3, \| \vec{v} \| = 2, \| \vec{w} \| = 1$. Then evaluate the following expressions :

(a) $< \vec{u} - \vec{v} - 2\vec{w}, 4\vec{u} + \vec{v} >$ (b) $\| 2\vec{w} - \vec{u} \|$.

Solution : (a) We use properties of inner product to simplify

$$< \vec{u} - \vec{v} - 2\vec{w}, 4\vec{u} + \vec{v} > = < \vec{u}, 4\vec{u} + \vec{v} > - < \vec{v}, 4\vec{u} + \vec{v} > - < 2\vec{w}, 4\vec{u} + \vec{v} >$$

$$= < \vec{u}, 4\vec{u} > + < \vec{u}, \vec{v} > - < \vec{v}, 4\vec{u} > - < \vec{v}, \vec{v} > - < 2\vec{w}, 4\vec{u} > - < 2\vec{w}, \vec{v} >$$

$$= 4 < \vec{u}, \vec{u} > + < \vec{u}, \vec{v} > - 4 < \vec{v}, \vec{u} > - < \vec{v}, \vec{v} > - 8 < \vec{w}, \vec{u} > - 2 < \vec{w}, \vec{v} >$$

$$= 4 \| \vec{u} \|^2 - 3 < \vec{u}, \vec{v} > - \| \vec{v} \|^2 - 8 < \vec{w}, \vec{u} > - 2 < \vec{w}, \vec{v} >$$

$$= 4 \times 9 - 3 \times (-1) - 4 - 8 \times 2 - 2 \times 3 = 36 + 3 - 4 - 16 - 6$$

$$= 13.$$

(b)
$$\| 2\vec{w} - \vec{u} \|^2 = < 2\vec{w} - \vec{u}, 2\vec{w} - \vec{u} >$$

$$= < 2\vec{w}, 2\vec{w} - \vec{u} > - < \vec{u}, 2\vec{w} - \vec{u} >$$

$$= < 2\vec{w}, 2\vec{w} > - < 2\vec{w}, \vec{u} > - < \vec{u}, 2\vec{w} > + < \vec{u}, \vec{u} >$$

$$= 4 < \vec{w}, \vec{w} > - 2 < \vec{w}, \vec{u} > - 2 < \vec{u}, \vec{w} > + < \vec{u}, \vec{u} >$$

$$= 4 \| \vec{w} \|^2 - 4 < \vec{u}, \vec{w} > + \| \vec{u} \|^2$$

$$= 4 - 4 \times 2 + 9$$

$$= 5$$

$$\therefore \quad \| 2\vec{w} - \vec{u} \| = \sqrt{5}.$$

Example 6.18 : Let P_2 be a vector space with inner product

$$< p, q > = \int_{-1}^{1} p(x)\, q(x)\, dx.$$

Find (a) $\| p \|$ for $p = 1 - x$, (b) $d(p, q)$, if $p = x$, and $q = x^2$.

Solution : (a) $\| p \| = < p, p >^{1/2}$

$$= \left[\int_{-1}^{1} (1 - x)^2 \, dx \right]^{1/2} = \left[\left(\frac{-(1-x)^3}{3} \right)_{-1}^{1} \right]^{1/2} = 2\sqrt{\frac{2}{3}}.$$

(b) $d(p, q) = \| p - q \| = < p - q, p - q >^{1/2}$

$$= \left[\int_{-1}^{1} (x - x^2)^2 \, dx \right]^{1/2} \qquad \because p - q = x - x^2$$

$$= \left[\int_{-1}^{1} (x^2 - 2x^3 + x^4) \, dx \right]^{1/2} = \left[\left(\frac{x^3}{3} - \frac{2x^4}{4} + \frac{x^5}{5} \right)_{-1}^{1} \right]^{1/2}$$

$$= \left[\left(\frac{1}{3} - \frac{2}{4} + \frac{1}{5} \right) - \left(-\frac{1}{3} - \frac{2}{4} - \frac{1}{5} \right) \right]^{1/2}$$

$$= \left(\frac{2}{3} + \frac{2}{5} \right)^{1/2} = \frac{4}{\sqrt{15}}.$$

Example 6.19 : Show that the following identities hold for vectors in any inner product space :

(a) $\| \vec{u} + \vec{v} \|^2 + \| \vec{u} - \vec{v} \|^2 = 2 \| \vec{u} \|^2 + 2 \| \vec{v} \|^2$

(b) $\| \vec{u} + \vec{v} \|^2 - \| \vec{u} - \vec{v} \|^2 = 4 < \vec{u}, \vec{v} >$

Solution : (a) We have

$$\begin{aligned} \| \vec{u} + \vec{v} \|^2 + \| \vec{u} - \vec{v} \|^2 &= < \vec{u} + \vec{v}, \vec{u} + \vec{v} > + < \vec{u} - \vec{v}, \vec{u} - \vec{v} > \\ &= < \vec{u}, \vec{u} + \vec{v} > + < \vec{v}, \vec{u} + \vec{v} > \\ &\quad + < \vec{u}, \vec{u} - \vec{v} > - < \vec{v}, \vec{u} - \vec{v} > \\ &= < \vec{u}, \vec{u} > + < \vec{u}, \vec{v} > + < \vec{v}, \vec{u} > + < \vec{v}, \vec{v} > \\ &\quad + < \vec{u}, \vec{u} > - < \vec{u}, \vec{v} > - < \vec{v}, \vec{u} > + < \vec{v}, \vec{v} > \\ &= 2 < \vec{u}, \vec{u} > + 2 < \vec{v}, \vec{v} > \\ &= 2 \| \vec{u} \|^2 + 2 \| \vec{v} \|^2 \end{aligned}$$

(b) Again

$$\begin{aligned} \| \vec{u} + \vec{v} \|^2 - \| \vec{u} - \vec{v} \|^2 &= < \vec{u} + \vec{v}, \vec{u} + \vec{v} > - < \vec{u} - \vec{v}, \vec{u} - \vec{v} > \\ &= < \vec{u}, \vec{u} + \vec{v} > + < \vec{v}, \vec{u} + \vec{v} > - < \vec{u}, \vec{u} - \vec{v} > \\ &\quad + < \vec{v}, \vec{u} - \vec{v} > \\ &= < \vec{u}, \vec{u} > + < \vec{u}, \vec{v} > + < \vec{v}, \vec{u} > + < \vec{v}, \vec{v} > \\ &\quad - < \vec{u}, \vec{u} > + < \vec{u}, \vec{v} > + < \vec{v}, \vec{u} > - < \vec{u}, \vec{v} > \\ &= 4 < \vec{u}, \vec{v} >. \end{aligned}$$

Example 6.20 : In each part use the inner product $< f, g > = \int_0^1 f(x)\, g(x)\, dx$ and compute $< f, g >$ for the vectors $f = f(x)$ and $g = g(x)$ in the function space of continuous functions defined on [0, 1].

(a) $f(x) = \cos(2\pi x)$, $g(x) = \sin 2\pi x$ (b) $f(x) = x$, $g(x) = e^x$.

Solution :

(a) $$\begin{aligned} < f, g > &= \int_0^1 \cos 2\pi x \sin 2\pi x \, dx = \frac{1}{2} \int_0^1 2 \sin 2\pi x \cos 2\pi x \, dx \\ &= \frac{1}{2} \int_0^1 \sin(4\pi x)\, dx = \frac{1}{2} \left(-\frac{\cos(4\pi x)}{4\pi} \right)_0^1 \\ &= \frac{1}{8\pi} (-\cos 4\pi + \cos 0) = \frac{1}{8\pi} (-1 + 1) = 0. \end{aligned}$$

(b) $$< f, g > = \int_0^1 x\, e^x\, dx$$

$$= (x\, e^x)_0^1 - \int_0^1 e^x\, dx, \qquad \text{integrating by parts}$$

$$= (e - 0) - (e^x)_0^1 = e - (e^1 - e^0)$$

$$= 1. \qquad \because\ e^0 = 1.$$

Example 6.21 : Let $M_{2 \times 2}(\mathbb{R})$ have inner product as in example (9). Find d (A, B), where, $A = \begin{pmatrix} -2 & 1 \\ 4 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} -5 & 6 \\ 1 & 2 \end{pmatrix}$.

Solution : $d(A, B) = \| A - B \|$

$$= \left\| \begin{pmatrix} 3 & 3 \\ -5 & -2 \end{pmatrix} \right\|$$

$$= \sqrt{3^2 + 3^2 + (-5)^2 + (-2)^2}$$

$$= \sqrt{47}\ .$$

Example 6.22 : Let $\vec{u} = (-1, 1, 2)$ and $\vec{v} = (2, -1, 2)$,

then find $< 2\vec{u} - 3\vec{v}, 3\vec{u} + \vec{v} >$.

Solution : Consider,

$$< 2\vec{u} - 3\vec{v}, 3\vec{u} + \vec{v} > = < 2\vec{u}, 3\vec{u} + \vec{v} > - < 3\vec{v}, 3\vec{u} + \vec{v} >$$

$$= < 2\vec{u}, 3\vec{u} > + < 2\vec{u}, \vec{v} > - < 3\vec{v}, 3\vec{u} > - < 3\vec{v}, \vec{v} >$$

$$= 6 < \vec{u}, \vec{u} > + 2 < \vec{u}, \vec{v} > - 9 < \vec{v}, \vec{u} > - 3 < \vec{v}, \vec{v} >$$

$$= 6 \| \vec{u} \|^2 - 7 < \vec{u}, \vec{v} > - 3 \| \vec{v} \|^2$$

As $\| \vec{u} \|^2 = (-1)^2 + (1)^2 + (2)^2 = 6$, $\| \vec{v} \|^2 = (2)^2 + (-1)^2 + (2)^2 = 9$

and $< \vec{u}, \vec{v} > = (-1) \times 2 + 1 \times (-1) + 2 \times 2 = 1$.

Substituting these values, we obtain

$$< 2\vec{u} - 3\vec{v}, 3\vec{u} + \vec{v} > = 6 \times 6 - 7 \times 1 - 3 \times 9$$

$$= 2.$$

Alternatively :

$$2\vec{u} - 3\vec{v} = (-2, 2, 4) + (-6, 3, -6) = (-8, 5, -2)$$

and $$3\vec{u} + \vec{v} = (-3, 3, 6) + (2, -1, 2) = (-1, 2, 8)$$

$$\therefore < 2\vec{u} - 3\vec{v}, 3\vec{u} + \vec{v} > = (-8) \times (-1) + 5 \times 2 + (-2) \times 8$$

$$= 2.$$

Exercise 6.1

1. In each part compute the Euclidean norm of the vectors :
   (a) (2, –3) (b) (1, 2, 3) (c) (4, –3, 0, 12) (d) (–2, 1, –3, 1, 4)
2. Find the Euclidean inner product $u \cdot v$.
   (a) $\vec{u} = (-2, 5)$, $\vec{v} = (4, -5)$
   (b) $\vec{u} = (8, 4, 2)$, $\vec{v} = (0, -1, 3)$
   (c) $\vec{u} = (-5, 4, 3, 1)$, $\vec{v} = (-4, -3, 2, 2)$
   (d) $\vec{u} = (-1, 1, 0, 4, -3)$, $\vec{v} = (2, -2, 0, 2, -1)$
3. Find all scalars k such that $\| ku \| = 7$, where $\vec{u} = (-2, 0, 3, 6)$.
4. Let $< \vec{u}, \vec{v} >$ be the Euclidean inner product on $\mathbb{R}^2$, and $\vec{u} = (-2, 3, 1)$, $\vec{v} = (5, -4, 0)$,
   $\vec{w} = (6, -1, -2)$ and $k = -3$. Verify each of the following parts :
   (a) $< \vec{u}, \vec{v} > = < \vec{v}, \vec{u} >$
   (b) $< \vec{u} + \vec{v}, \vec{w} > = < \vec{u}, \vec{w} > + < \vec{v}, \vec{w} >$
   (c) $< \vec{u}, \vec{v} + \vec{w} > = < \vec{u}, \vec{v} > + < \vec{u}, \vec{w} >$
   (d) $< \vec{0}, \vec{u} > = < \vec{u}, \vec{0} > = 0$
   (e) $< k\vec{u}, \vec{v} > = k < \vec{u}, \vec{v} > = < \vec{u}, k\vec{v} >$
5. Repeat exercise (4) using the inner product
   $< \vec{u}, \vec{v} > = 3 u_1v_1 + 2 u_2v_2 + u_3v_3$.
6. (a) Show that $< \vec{u}, \vec{v} > = 10 u_1v_1 + 14 u_1v_2 + 14 u_2v_1 + 20 u_2v_2$ is the inner product on $\mathbb{R}^2$ generated by the matrix
   $$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$
   using the formula (**) of example (6).
   (b) Use the inner product in (a) to compute $< \vec{u}, \vec{v} >$, if $\vec{u} = (0, -2)$ and $\vec{v} = (3, 1)$.
7. Let $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$. In each part, the given expression is an inner product on $\mathbb{R}^2$. Find a matrix that generates it.
   (a) $< \vec{u}, \vec{v} > = -u_1v_1 - 2 u_1 v_2 + u_2v_1 - 2u_2v_2$ (b) $< \vec{u}, \vec{v} > = 4 u_1v_1 + 4 u_2 v_2$.

8. Let $\vec{u} = (u_1, u_2)$ and $\vec{v} = (v_1, v_2)$. Show that the following are inner products on $\mathbb{R}^2$, by verifying that the inner product axioms hold.

   (a) $< \vec{u}, \vec{v} > = 2\,u_1v_1 + 3\,u_2v_2$

   (b) $< \vec{u}, \vec{v} > = 4\,u_1v_1 + u_2v_1 + u_1v_2 + 4\,u_2v_2$.

9. Let $\vec{u} = (u_1, u_2, u_3)$ and $\vec{v} = (v_1, v_2, v_3)$. Determine which of the following are inner products on $\mathbb{R}^3$. For those that are not, list the axioms that do not hold.

   (a) $< \vec{u}, \vec{v} > = u_1\,v_1 + u_3\,v_3$ (b) $< \vec{u}, \vec{v} > = u_1^2\,v_1^2 + u_2^2\,v_2^2 + u_3^2\,v_3^2$

   (c) $< \vec{u}, \vec{v} > = 2\,u_1v_1 + u_2v_2 + 4\,u_3v_3$ (d) $< \vec{u}, \vec{v} > = u_1v_1 - u_2v_2 + u_3v_3$

10. Suppose that $\vec{u}$, $\vec{v}$, and $\vec{w}$ are vectors such that $< \vec{u}, \vec{v} > = 2$, $< \vec{v}, \vec{w} > = -3$, $< \vec{u}, \vec{w} > = 5$, $\| \vec{u} \| = 1$, $\| \vec{v} \| = 2$, $\| \vec{w} \| = 3$. Evaluate each of the following expressions :

    (a) $< \vec{u} + \vec{v}, \vec{v} + \vec{w} >$ (b) $< 2\vec{v} - \vec{w}, 3\vec{u} + 2\vec{w} >$

    (c) $< \vec{u} - \vec{v} - 2\vec{w}, 4\vec{u} + \vec{v} >$ (d) $\| \vec{u} + \vec{v} \|$

    (e) $\| (2\vec{w} - \vec{v}) \|$ (f) $\| \vec{u} - 2\vec{v} + 4\vec{w} \|$

11. Let P_2 be the vector space with inner product

$$< p, q > = \int_{-1}^{1} p(x)\,q(x)\,dx$$

    (a) Find $\| p \|$ for $p = 1$, $p = 1 - x$, and $p = x^2$

    (b) Find $< p, q >$, if $p = 1 - x$, $q = x$.

12. Let p and q be polynomials in P_2. Show that

$$< p, q > = p(0)\,q(0) + p\left(\frac{1}{2}\right)\,q\left(\frac{1}{2}\right) + p(1)\,q(1)$$

    is an inner product on P_2.

13. If $< \vec{u}, \vec{v} >$ is the Euclidean inner product on $\mathbb{R}^n$, and if A is an $n \times n$ matrix, then show that $< \vec{u}, A\vec{v} > = < A^t\vec{u}, \vec{v} >$.

14. Let $\vec{u} = (2, i, 3)$ and $\vec{v} = (1 + 2i, 2i, 1)$ be vectors in $\mathbb{C}^3$, then find

    $< \vec{u}, \vec{v} >$, $< \vec{v}, \vec{u} >$, $\| \vec{u} \|$ and $\| \vec{v} \|$.

15. Compute (a) $< \vec{u}, \vec{v} >$, (b) $< \vec{v}, \vec{u} >$, (c) $\| \vec{u} \|, \| \vec{v} \|$, (d) $< \alpha \vec{u}, \beta \vec{v} >$, (e) $\| \vec{u} \ \vec{v} \|$, where

(i) $\vec{u} = (i, 1, -1)$, $\vec{v} = (1, -i, 1)$, $\alpha = 1$, $\beta = 2i$.

(ii) $\vec{u} = (1, -1, 2)$, $\vec{v} = (1 - i, 1 + i, 0)$, $\alpha = 1 + i$, $\beta = 2$.

Answers (6.1)

1. (a) $\sqrt{13}$ (b) $\sqrt{15}$ (c) 13 (d) $\sqrt{31}$.
2. (a) – 33 (b) 2 (c) 16 (d) 7
3. ± 1.
6. (b) – 124
7. (a) $\begin{bmatrix} 1 & -2 \\ 1 & 0 \end{bmatrix}$, (b) $\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$
9. (a) No, Axiom (4) fails (b) No, Axioms (2) and (3) fail.
   (c) Yes (d) No, Axiom (4) fails.
10. (a) 8 (b) – 33 (c) – 40 (d) 3 (e) $2\sqrt{13}$, (f) $\sqrt{241}$.
11. (a) $\sqrt{2}$, $2\sqrt{2/3}$, $\sqrt{\frac{2}{5}}$ (b) $-\frac{4}{6}$.
14. $< \vec{u}, \vec{v} > = 7 - 4i$, $< \vec{v}, \vec{u} > = 7 + 4i$, $\| \vec{u} \| = \sqrt{14}$, $\| \vec{v} \| = \sqrt{10}$.
15. (i) (a) $-1 + 2i$, (b) $-1 - 2i$, (c) $\sqrt{3}, \sqrt{3}$, (d) $4 + 2i$, (e) $2\sqrt{2}$.
    (ii) (a) $2i$, (b) $-2i$, (c) $\sqrt{6}, 2$, (d) $2(-1 + i)$, (e) $\sqrt{10}$

6.4 Angle and Orthogonality in an Inner Product Space

In this section we shall define the notion of an angle between two vectors in an inner product space, and we use this concept to obtain important properties of vectors in inner product spaces analogues to Euclidean geometry.

We know that, if $\vec{u}$ and $\vec{v}$ are non-zero vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ and θ is the angle between them, then by the relation (1) in section 6.2,

$$\vec{u} \cdot \vec{v} = \| \vec{u} \| \| \vec{v} \| \cos \theta, \text{ in } \mathbb{R}^n \ | \vec{u} | = \| \vec{u} \|$$

This relation can be written as

$$\cos \theta = \frac{\vec{u} \cdot \vec{v}}{\| \vec{u} \| \| \vec{v} \|}$$

Squaring both sides in first relation and using the fact that $\|\vec{u}\|^2 = \vec{u} \cdot \vec{u}$, $\|\vec{v}\|^2 = \vec{v} \cdot \vec{v}$ and $\cos^2 \theta \leq 1$, we get

$$(\vec{u} \cdot \vec{v})^2 \leq (\vec{u} \cdot \vec{u})(\vec{v} \cdot \vec{v})$$

This inequality can be generalized to any general vector space V over $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). We first prove here **Schwartz inequality** which is analogous to the above inequality.

Theorem 3 : (Schwartz Inequality)

If $\vec{u}$ and $\vec{v}$ are two vectors in an inner product space, then

$$[<\vec{u}, \vec{v}><\vec{u}, \vec{v}>]^{1/2} = |<\vec{u}, \vec{v}>| \leq \|\vec{u}\| \|\vec{v}\| \quad \ldots (1)$$

Proof : We observe that if $\vec{v} = \vec{0}$, then the inequality holds trivially. So suppose $\vec{v} \neq \vec{0}$, and let us define vector $\vec{w} = \dfrac{\vec{v}}{\|\vec{v}\|}$ ($\vec{w}$ is normalized vector $\vec{v}$).

We take a scalar $\alpha = <\vec{u}, \vec{w}>$

$$= <\vec{u}, \frac{\vec{v}}{\|\vec{v}\|}>$$

$$= \frac{1}{\|\vec{v}\|} <\vec{u}, \vec{v}>$$

It is clear that $<\vec{w}, \vec{w}> = 1$, since it is unit vector. Now using the axiom of non-negativity, we have

$$0 \leq <\vec{u} - \alpha\vec{w}, \vec{u} - \alpha\vec{w}> = <\vec{u}, \vec{u} - \alpha\vec{w}> - \alpha <\vec{w}, \vec{u} - \alpha\vec{w}> \text{ Using axiom (3)}$$

$$= <\vec{u}, \vec{u}> - \bar{\alpha} <\vec{u}, \vec{w}> - \alpha <\vec{w}, \vec{u}> - \bar{\alpha}\alpha <\vec{w}, \vec{w}>$$

$$= \|\vec{u}\|^2 - \bar{\alpha}\alpha - \alpha\bar{\alpha} + \alpha\bar{\alpha}$$

$$\because <\vec{w}, \vec{w}> = 1 \text{ and } <\vec{w}, \vec{u}> = \overline{<\vec{u}, \vec{w}>}$$

$$= \|\vec{u}\|^2 - |\alpha|^2$$

Now replacing α by $<\vec{u}, \vec{w}>$, we obtain

$$0 \leq \|\vec{u}\|^2 - |<\vec{u}, \vec{w}>|^2$$

or $$|<\vec{u}, \vec{w}>|^2 \leq \|\vec{u}\|^2$$

or $$\left| < \vec{u}, \frac{\vec{v}}{\|\vec{v}\|} > \right|^2 \leq \|\vec{u}\|^2 \qquad \because \vec{w} = \frac{\vec{v}}{\|\vec{v}\|}$$

or $$\frac{1}{\|\vec{v}\|^2} |< \vec{u}, \vec{v} >|^2 \leq \|\vec{u}\|^2$$

or $$|< \vec{u}, \vec{v} >|^2 \leq \|\vec{u}\|^2 \|\vec{v}\|^2$$

Taking positive square root on both sides, we get the required inequality.

$$|< \vec{u}, \vec{v} >| \leq \|\vec{u}\| \|\vec{v}\|$$

An application of theorem (3) gives the following :

Corollary : (Schwartz inequality) :

If $\vec{u}$ and $\vec{v}$ are two vectors in an Euclidean space $\mathbb{R}^n$, then,

$$-1 \leq \frac{< \vec{u}, \vec{v} >}{\|\vec{u}\| \|\vec{v}\|} \leq 1 \qquad \ldots (2)$$

Proof : Using Schwartz inequality proved in theorem (3), we have

$$|< \vec{u}, \vec{v} >| \leq \|\vec{u}\| \|\vec{v}\|$$

This inequality can be written as

$$\frac{1}{\|\vec{u}\| \|\vec{v}\|} |< \vec{u}, \vec{v} >| \leq 1$$

or $$\left| \frac{1}{\|\vec{u}\| \|\vec{v}\|} < \vec{u}, \vec{v} > \right| \leq 1$$

or $$-1 \leq \frac{1}{\|\vec{u}\| \|\vec{v}\|} < \vec{u}, \vec{v} > \leq 1 \quad \text{property of absolute value.}$$

or $$-1 \leq \frac{< \vec{u}, \vec{v} >}{\|\vec{u}\| \|\vec{v}\|} \leq 1$$

This proves the corollary.

Remark : If θ is the angle whose radian measure varies from 0 to π, then $\cos\theta$ assumes every value between -1 and 1 inclusive exactly once. Therefore, we see that the relation (2) determines unique angle θ such that

$$\cos\theta = \frac{< \vec{u}, \vec{v} >}{\|\vec{u}\| \|\vec{v}\|} \text{ and } 0 \leq \theta \leq \pi$$

θ is the angle between $\vec{u}$ and $\vec{v}$. This is possible only in **real vector** spaces. In complex vector spaces it is not possible as $< \vec{u}, \vec{v} >$ in that case is a complex number, hence inequality (2) can not hold.

SOLVED EXAMPLES

Example 6.23 : Let $\mathbb{R}^4$ have the Euclidean inner product. Find the cosine of the angle between the vectors $\vec{u} = (-1, 2, 3, 4)$ and $\vec{v} = (4, 1, 2, 1)$.

Solution : We have, $\| \vec{u} \| = \sqrt{30}$; $\| \vec{v} \| = \sqrt{22}$ and $< \vec{u}, \vec{v} > = 8$.

So that $$\cos\theta = \frac{< \vec{u}, \vec{v} >}{\| \vec{u} \| \| \vec{v} \|}$$

$$= \frac{8}{\sqrt{30} \cdot \sqrt{22}} = \frac{4}{\sqrt{165}}$$

Example 6.24 : Let $M_{2\times 2}(\mathbb{R})$ be a vector space with inner product. Show that the angle between the matrices $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$ is $\frac{\pi}{2}$.

Solution : We have $< A, B > = 1 \times 1 + 1 \times 0 + 0 \times 2 + 1 \times (-1) = 0$.

$\therefore$ $$\cos\theta = \frac{< A, B >}{\| A \| \| B \|} = \frac{0}{\| A \| \| B \|} = 0.$$

Thus, $$\theta = \frac{\pi}{2}.$$

Example 6.25 : If $\vec{u}$ and $\vec{v}$ are non-zero vectors in an inner product space, and if θ is the angle between $\vec{u}$ and $\vec{v}$, then $\theta = \pi/2$ if and only if $< \vec{u}, \vec{v} > = 0$.

Solution : We know

$$\cos\theta = \frac{< \vec{u}, \vec{v} >}{\| \vec{u} \| \| \vec{v} \|}$$

If $< \vec{u}, \vec{v} > = 0$, then $\cos\theta = 0$, thus $\theta = \frac{\pi}{2}$. On the other hand, if $\theta = \frac{\pi}{2}$, then $\cos\theta = 0$.

Therefore $$\frac{< \vec{u}, \vec{v} >}{\| \vec{u} \| \| \vec{v} \|} = 0 \Rightarrow < \vec{u}, \vec{v} > = 0.$$

Definition : In an inner product space V, two vectors $\vec{u}$ and $\vec{v}$ are called *orthogonal* if $< \vec{u}, \vec{v} > = 0$. If $\vec{u}$ is orthogonal to each vector in a set W, we say that $\vec{u}$ *is orthogonal to W.*

Example 6.26 : Let P_2 be a vector space with inner product

$$< p, q > = \int_{-1}^{1} p(x)\, q(x)\, dx.$$

If p (x) = 1 and q (x) = x, show that p and q are orthogonal in P_2.

Solution : We have

$$< p, q > = \int_{-1}^{1} p(x)\, q(x)\, dx = \int_{-1}^{1} 1 \cdot x\, dx = \left(\frac{x^2}{2}\right)_{-1}^{+1} = 0$$

Thus, p and q are orthogonal in P_2.

Theorem 4 : (Generalized Theorem of Pythagorus).

If $\vec{u}$ and $\vec{v}$ are orthogonal vectors in an inner product space, then

$$\| \vec{u} + \vec{v} \|^2 = \| \vec{u} \|^2 + \| \vec{v} \|^2$$

Proof : As $\vec{u}$ and $\vec{v}$ are orthogonal, $< \vec{u}, \vec{v} > = 0$.

Consider $\| \vec{u} + \vec{v} \|^2 = < \vec{u} + \vec{v}, \vec{u} + \vec{v} >$

$$= \| \vec{u} \|^2 + 2 < \vec{u}, \vec{v} > + \| \vec{v} \|^2$$

$$= \| \vec{u} \|^2 + \| \vec{v} \|^2 \qquad \because < \vec{u}, \vec{v} > = 0.$$

Note : In $\mathbb{R}^2$ or $\mathbb{R}^3$ with Euclidean inner product, this theorem reduces to usual Pythagorus theorem.

SOLVED EXAMPLES

Example 6.27 : Let P_2 be the vector space with inner product $<p, q> = \int_{-1}^{1} p(x)\, q(x)\, dx$. If p (x) = 1 and q (x) = x, show that p and q are orthogonal. Further verify Pythagorus theorem.

Solution : In example (5), we showed that p and q are orthogonal.

We have

$$\| p \| = < p, p >^{1/2}$$

$$= \left[\int_{-1}^{1} (p(x))^2\, dx \right]^{1/2} = \left[\int_{-1}^{1} 1\, dx \right]^{1/2} = \sqrt{2}\,.$$

$$\| q \| = < q, q >^{1/2} = \left[\int_{-1}^{1} (q(x))^2 \, dx \right]^{1/2}$$

$$= \left[\int_{-1}^{1} x^2 \, dx \right]^{1/2}$$

$$= \left[\left(\frac{x^3}{3} \right)_{-1}^{1} \right]^{1/2} = \sqrt{\frac{2}{3}} \, .$$

$$\| p + q \| = < p + q, p + q >^{1/2} = \left[\int_{-1}^{1} (1 + x)^2 \, dx \right]^{1/2}$$

$$= \left[\left\{ \frac{(1 + x)^3}{3} \right\}_{-1}^{1} \right]^{1/2}$$

$$= \left(\frac{8}{3} - 0 \right)^{1/2} = 2\sqrt{\frac{2}{3}} \, .$$

Now, $\| p \|^2 + \| q \|^2 = (\sqrt{2})^2 + \left(1\sqrt{\frac{2}{3}} \right)^2 = 2 + \frac{2}{3} = \frac{8}{3}$

and $\| p + q \|^2 = \left(2\sqrt{\frac{2}{3}} \right)^2 = \frac{8}{3}$.

Thus $\| p + q \|^2 = \| p \|^2 + \| q \|^2$.

Example 6.28 : Let $\mathbb{R}^4$ have Euclidean inner product. Find two vectors of norm 1 that are orthogonal to the three vectors $\vec{u} = (2, 1, -4, 0)$, $\vec{v} = (-1, -1, 2, 2)$, and $\vec{w} = (3, 2, 5, 4)$.

Solution : Let $\vec{X} = (x_1, x_2, x_3, x_4)$ be a vector orthogonal to the three vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$. Then

$$\vec{X} \cdot \vec{u} = 0, \vec{X} \cdot \vec{v} = 0 \text{ and } \vec{X} \cdot \vec{w} = 0$$

This gives

$$2x_1 + x_2 - 4x_3 = 0 \quad \text{... (i)}$$

$$-x_1 - x_2 + 2x_3 + 2x_4 = 0 \quad \text{... (ii)}$$

$$3x_1 + 2x_2 + 5x_3 + 4x_4 = 0 \quad \text{... (iii)}$$

This is homogeneous system of linear equations in x_1, x_2, x_3 and x_4.

Solving this system by Gauss-elimination method (students are requested to elaborate the calculations), we get a general solution as :

$$x_1 = -\frac{34}{11} d, \; x_2 = 4d, \; x_3 = -\frac{6}{11} d \text{ and } x_4 = d, \; d \in \mathbb{R}.$$

Taking d = 1, we get

$\vec{X} = \left(-\frac{34}{11}, 4, -\frac{6}{11}, 1\right)$ is the vector orthogonal to $\vec{u}$, $\vec{v}$ and $\vec{w}$. For norm to be 1, we normalize $\vec{X}$. For this, we have

$$\| \vec{X} \| = \sqrt{\left(-\frac{34}{11}\right)^2 + 16 + \left(-\frac{6}{11}\right)^2 + 1}$$

$$= \sqrt{\frac{3249}{121}} = \frac{57}{11}.$$

$\therefore$ The two vectors orthogonal to the three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ of norm 1 are given by $\pm \frac{11}{57}\left(-\frac{34}{11}, 4, -\frac{6}{11}, 1\right)$ or $\pm \frac{1}{57}(-34, 44, -6, 11)$

Example 6.29 : Using Euclidean inner product on $\mathbb{R}^3$, show that $\vec{u} = (-3, 1, 0)$ and $\vec{v} = (2, -1, 3)$ satisfy Cauchy-Schwarz inequality.

Solution : We have

$$< \vec{u}, \vec{v} > = (-3)(2) + (1)(-1) + (0)(3)$$

$$= -7$$

$$\| \vec{u} \| = \sqrt{10}, \quad \| \vec{v} \| = \sqrt{14}$$

Now, the Cauchy-Schwarz inequality is

$$< \vec{u}, \vec{v} >^2 \leq \| \vec{u} \|^2 \| \vec{v} \|^2 \quad \text{... (i)}$$

L. H. S. $< \vec{u}, \vec{v} >^2 = (-7)^2 = 49$

and R.H.S. $= \| \vec{u} \|^2 \| \vec{v} \|^2$

$= (\sqrt{10})^2 (\sqrt{14})^2 = 140$

$\therefore$ L.H.S. $<$ R.H.S.

Hence (i) is satisfied.

Example 6.30 : Let V be an inner product space. If $\vec{u}$ and $\vec{v}$ are orthogonal vectors in V such that $\| \vec{u} \| = 1$, and $\| \vec{v} \| = 1$, then show that

(a) $\| \vec{u} + \vec{v} \|^2 + \| \vec{u} - \vec{v} \|^2 = 4$ (b) $\| \vec{u} + \vec{v} \| = \sqrt{2}$.

Solution : (a) We know that, in any inner product space,

$$\| \vec{u} + \vec{v} \|^2 + \| \vec{u} - \vec{v} \|^2 = 2 \| \vec{u} \|^2 + 2 \| \vec{v} \|^2$$

$\therefore$ $\| \vec{u} + \vec{v} \|^2 + \| \vec{u} - \vec{v} \|^2 = 2 \times 1 + 2 \times 1 = 4.$

(b) We have $\| \vec{u} + \vec{v} \|^2 = < \vec{u} + \vec{v}, \vec{u} + \vec{v} >^{1/2}$

$= [\| \vec{u} \|^2 + 2 < \vec{u}, \vec{v} > + \| \vec{v} \|^2]^{1/2}$

$= (1 + 2 \times 0 + 1)^{1/2}, \qquad \because < \vec{u}, \vec{v} > = 0$

$= \sqrt{2}$.

Example 6.31 : Let $\{\vec{v_1}, \vec{v_2}, ... , \vec{v_n}\}$ be a basis for an inner product space V. Show that the zero vector is the only vector in V that is orthogonal to all of the basis vectors.

Solution : Suppose $\vec{u}$ is a vector orthogonal to all the basis vectors $\vec{v_1}, \vec{v_2}, ... , \vec{v_n}$.

Then $< \vec{u}, \vec{v_i} > = 0$, for each i = 1, 2, ... , n.

As $\{\vec{v_1}, \vec{v_2}, ... , \vec{v_n}\}$ is a basis for V, $\vec{u}$ is a linear combination of the basis vectors, that is, $\vec{u} = \alpha_1\vec{v_1} + \alpha_2\vec{v_2} + ... + \alpha_n\vec{v_n}$, where $\alpha_1, \alpha_2, ... , \alpha_n$ are the scalars.

Now $< \vec{u}, \vec{u} > = < \vec{u}, \alpha_1\vec{v_1} + \alpha_2\vec{v_2} + ... + \alpha_n\vec{v_n} >$

$= \alpha_1 < \vec{u}, \vec{v_1} > + \alpha_2 < \vec{u}, \vec{v_2} > +.. + \alpha_n < \vec{u}, \vec{v_n} >$

$= \alpha_1 \cdot 0 + \alpha_2 \cdot 0 + ... + \alpha_n \cdot 0$

$= 0.$

i.e. $< \vec{u}, \vec{u} > = 0$ hence $\vec{u} = \vec{0}$, by axiom (4).

The result follows.

Example 6.32 : Let V be an inner product space and $\vec{v}$ be any vector in V. Show that the set W, of vectors in V that are orthogonal to $\vec{v}$ forms a subspace of V.

Solution : To show that W is a subspace of V, we need to show that W is closed under addition and scalar multiplication.

First note that for $\vec{0}$ in V $< \vec{0}, \vec{v} > = 0$, hence $\vec{0}$ is in W. Let $\vec{w_1}, \vec{w_2}$ be any two vectors in W, then $< \vec{v}, \vec{w_1} > = < \vec{v}, \vec{w_2} > = 0$. For any scalars α_1, α_2, consider

$< \vec{v}, \alpha_1\vec{w_1} + \alpha_2 \vec{w_2} >$

$= \bar{\alpha}_1 < \vec{v}, \vec{w_1} > + \bar{\alpha}_2 < \vec{v}, \vec{w_2} >$

$= \bar{\alpha}_1 \cdot 0 + \bar{\alpha}_2.$

$= 0.$

This shows that $\alpha_1\vec{w_1} + \alpha_2\vec{w_2}$ is in W. Therefore by alternative necessary and sufficient conditions for subspace, we see that W is a subspace of V.

Example 6.33 : For all real values of a, b and θ, show that

$[a \cos \theta + b \sin \theta]^2 \leq a^2 + b^2$.

Solution : For fixed real value of θ, (cos θ, sin θ) is a pair of real numbers and hence belongs to $\mathbb{R}^2$. Taking $\vec{u} = (a, b)$ and $\vec{v} = (\cos \theta, \sin \theta)$ and applying Cauchy-Schwarz inequality to $\vec{u}$ and $\vec{v}$ with Euclidean inner product, we have,

$$< \vec{u}, \vec{v} >^2 \leq \| \vec{u} \|^2 \| \vec{v} \|^2$$

$$(a \cos \theta + b \sin \theta)^2 \leq (a^2 + b^2)(\cos^2 \theta + \sin^2 \theta)$$

i.e. $(a \cos \theta + b \sin \theta)^2 \leq a^2 + b^2$. $\because \cos^2 \theta + \sin^2 \theta = 1$.

Example 6.34 : In any inner product space V and for $\vec{u}$ and $\vec{v}$ in V, show that

(a) $< \vec{u}, \vec{v} > = 0$ if and only if $\| \vec{u} + \vec{v} \| = \| \vec{u} - \vec{v} \|$

(b) $< \vec{u} + \vec{v}, \vec{u} - \vec{v} > = 0$ iff $\| \vec{u} \| = \| \vec{v} \|$

Solution : (a) Suppose $< \vec{u}, \vec{v} > = 0$. Then we have

$$\| \vec{u} + \vec{v} \|^2 = < \vec{u} + \vec{v}, \vec{u} + \vec{v} >$$

$$= \| \vec{u} \|^2 + 2 < \vec{u}, \vec{v} > + \| \vec{v} \|^2$$

$$= \| \vec{u} \|^2 - 2 < \vec{u}, \vec{v} > + \| \vec{v} \|^2 \qquad \because < \vec{u}, \vec{v} > = 0$$

$$= < \vec{u} - \vec{v}, \vec{u} - \vec{v} >$$

$$= \| \vec{u} - \vec{v} \|^2$$

Taking positive square root, we obtain

$$\| \vec{u} + \vec{v} \| = \| \vec{u} - \vec{v} \|.$$

Conversely, suppose $\| \vec{u} + \vec{v} \| = \| \vec{u} - \vec{v} \|$.

This implies, squaring $\| \vec{u} + \vec{v} \|^2 = \| \vec{u} - \vec{v} \|^2$

or $< \vec{u} + \vec{v}, \vec{u} + \vec{v} > = < \vec{u} - \vec{v}, \vec{u} - \vec{v} >$

or $\| \vec{u} \|^2 + 2 < \vec{u}, \vec{v} > + \| \vec{v} \|^2 = \| \vec{u} \|^2 - 2 < \vec{u}, \vec{v} > + \| \vec{v} \|^2$

or $4 < \vec{u}, \vec{v} > = 0$

or $< \vec{u}, \vec{v} > = 0$

(b) Suppose $< \vec{u} + \vec{v}, \vec{u} - \vec{v} > = 0$

Then $< \vec{u}, \vec{u} - \vec{v} > + < \vec{v}, \vec{u} - \vec{v} > = 0$

or $< \vec{u}, \vec{u} > - < \vec{u}, \vec{v} > + < \vec{v}, \vec{u} > - < \vec{v}, \vec{u} > = 0$

or $\| \vec{u} \|^2 - \| \vec{v} \|^2 = 0$

or $\| \vec{u} \|^2 = \| \vec{v} \|^2$

or $\| \vec{u} \| = \| \vec{v} \|$ (Taking positive square root)

Conversely, suppose $\| \vec{u} \| = \| \vec{v} \|$

Consider,

$$< \vec{u} + \vec{v}, \vec{u} - \vec{v} > = < \vec{u}, \vec{u} > - < \vec{u}, \vec{v} > + < \vec{v}, \vec{u} > - < \vec{v}, \vec{v} >$$

$$= \| \vec{u} \|^2 - \| \vec{v} \|^2$$

$$= 0, \qquad \because \| \vec{u} \| = \| \vec{v} \|$$

6.5 Orthonormal Bases; Gram-Schmidt Process

Definition : Let V be an inner product space. A basis B for V is said to be *orthogonal*, if all pairs of distinct vectors in B are orthogonal. An orthogonal basis in which each vector has norm 1 is called *orthonormal basis.*

Example 6.35 : Show that the set $B = \{ \vec{u_1}, \vec{u_2}, \vec{u_3} \}$, where $\vec{u_1} = (0, 1, 0)$, $\vec{u_2} = (1, 0, 1)$ and $\vec{u_3} = (1, 0, -1)$ is an orthogonal basis for $\mathbb{R}^3$ with Euclidean inner product. Is it orthonormal ?

Solution : Clearly,

$$< \vec{u_1}, \vec{u_2} > = 0 \times 1 + 0 \times 1 + 0 \times 1 = 0$$

$$< \vec{u_2}, \vec{u_3} > = 1 \times 1 + 0 \times 0 + 1 \times (-1) = 0$$

and $$< \vec{u_1}, \vec{u_3} > = 0 \times 1 + 1 \times 0 + 0 \times (-1) = 0$$

Thus, the vectors $\vec{u_1}$, $\vec{u_2}$ and $\vec{u_3}$ are pairwise orthogonal. Also, $\| \vec{u_1} \| = 1$, but $\| \vec{u_2} \| = \sqrt{2} = \| \vec{u_3} \|$. Hence B is not orthonormal. To see that B is basis, we need to check only that B is linearly independent. For this suppose

$$\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \alpha_3 \vec{u_3} = 0 \qquad \ldots (*)$$

or in terms of components,

$$(\alpha_2 + \alpha_3, \alpha_1, \alpha_2 - \alpha_3) = (0, 0, 0)$$

or equating the corresponding components,

$$\begin{aligned} \alpha_2 + \alpha_3 &= 0 \\ \alpha_1 &= 0 \\ \alpha_2 - \alpha_3 &= 0. \end{aligned}$$

Solving this system gives $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Thus the identity (*) holds only if $\alpha_1 = \alpha_2 = \alpha_3 = 0$, hence B is independent, therefore B is an orthogonal basis.

Remark 1 : The Euclidean norms of the vectors in the above example are $\| \vec{u_1} \| = 1$, $\| \vec{u_2} \| = \sqrt{2}$, $\| \vec{u_3} \| = \sqrt{2}$.

Given a vector $\vec{u} \neq \vec{0}$, we have a vector $\vec{v} = \dfrac{1}{\| \vec{u} \|} \cdot \vec{u}$, which has norm 1, since

$$\| \vec{v} \| = \left\| \frac{1}{\| \vec{u} \|} \cdot \vec{u} \right\| = \frac{1}{\| \vec{u} \|} \cdot \| \vec{u} \|, \text{ using } L_3.$$

$$= 1.$$

This process dividing a non-zero vector by its norm to obtain a vector of norm 1 is called normalizing $\vec{u}$.

We normalize the vectors $\vec{u_1}$, $\vec{u_2}$ and $\vec{u_3}$, and obtain

$$\vec{v_1} = \frac{\vec{u_1}}{\|\vec{u_1}\|+} = (0, 1, 0), \vec{v_2} = \frac{\vec{u_2}}{\|\vec{u_2}\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)$$

$$\vec{v_3} = \frac{1}{\| \vec{u_3} \|} \vec{u_3} = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}} \right).$$

The new set $S = \{ \vec{v_1}, \vec{v_2}, \vec{v_3} \}$ is orthonormal basis.

Problem : Check that each vector in S in remark (1) has norm 1 and distinct pairs of vectors in S are orthogonal.

Remark 2 : It is easy to check that the set $B = \{ \vec{i}, \vec{j}, \vec{k} \}$, where $\vec{i} = (1, 0, 0)$, $\vec{j} = (0, 1, 0)$ and $\vec{k} = (0, 0, 1)$ is orthonormal basis for $\mathbb{R}^3$, with Euclidean inner product.

Further, the standard basis $\vec{e_1} = (1, 0, 0, ... 0)$, $\vec{e_2} = (0, 1, 0, ... , 0)$, ... $\vec{e_n} = (0, 0, ... 1)$ for $\mathbb{R}^n$ with Euclidean inner product is orthonormal.

Given a vector in a vector space, it is pretty difficult to express it as a linear combination of the basis vectors. The following theorem shows that this is eased, if the basis is orthonormal basis.

Theorem 5 : If $S = \{\vec{v_1}, \vec{v_2}, .. , \vec{v_n}\}$ is an orthonormal basis for an inner product space V, and $\vec{u}$ is any vector in V, then

$$\vec{u} = < \vec{u}, \vec{v_1} > \vec{v_1} + < \vec{u}, \vec{v_2} > \vec{v_2} + ... + < \vec{u}, \vec{v_n} > \vec{v_n}.$$

Proof : Since $S = \{\vec{v_1}, \vec{v_2}, ..., \vec{v_n}\}$ is a basis, a vector $\vec{u}$ can be expressed in the form

$$\vec{u} = k_1\vec{v_1} + k_2\vec{v_2} + ... + k_n\vec{v_n}$$

Consider, for i = 1, 2, ... , n,

$$< \vec{u}, \vec{v_i} > = < k_1\vec{v_1} + k_2\vec{v_2} + ... + k_n\vec{v_n}, \vec{v_i} >$$

$$= k_1 < \vec{v_1}, \vec{v_i} > + k_2 < \vec{v_2}, \vec{v_i} > + ...+ k_n < \vec{v_n}, \vec{v_i} >$$

Since S is orthonormal set $< \vec{v_i}, \vec{v_i} > = \| \vec{v_i} \|^2 = 1$ and $< \vec{v_j}, \vec{v_i} > = 0$, if $j \neq i$

therefore, the above expression for $< \vec{u}, \vec{v_i} >$ simplifies to

$$< \vec{u}, \vec{v_i} > = k_i , \text{ for } i = 1, 2, ... , n.$$

Therefore,

$$\vec{u} = < \vec{u} , \vec{v_1} > \vec{v_1} + < \vec{u}, \vec{v_2} > \vec{v_2} + ... + < \vec{u}, \vec{v_n} > \vec{v_n}.$$

Remark : If $S = \{\vec{v_1}, \vec{v_2}, ... , \vec{v_n}\}$ is an orthonormal basis for an inner product space V, and $\vec{u}$ is any vector in V, then the above theorem shows that the co-ordinates of $\vec{u}$ relative to the basis S are $< \vec{u}, \vec{v_1} >, < \vec{u}, \vec{v_2} >, ... , < \vec{u}, \vec{v_n} >$ and

$$(\vec{u})_s = (< \vec{u}, \vec{v_1} >, < \vec{u}, \vec{v_2} >, ... , < \vec{u}, \vec{v_n} >)$$

is a co-ordinate vector of u relative to this basis.

SOLVED EXAMPLES

Example 6.36 : Let $S = \left\{(0, 1, 0) , \left(\frac{1}{\sqrt{2}} , 0, \frac{1}{\sqrt{2}}\right) , \left(\frac{1}{\sqrt{2}} , 0, -\frac{1}{\sqrt{2}}\right)\right\}$ be a basis for $\mathbb{R}^3$ with Euclidean inner product. Find co-ordinates and co-ordinate vector of $\vec{u} = (1, 2, 3)$ relative to the basis S.

Solution : From theorem 7, if $\vec{v_1} = (0, 1, 0)$, $\vec{v_2} = \left(\frac{1}{\sqrt{2}} , 0, \frac{1}{\sqrt{2}}\right)$ and $\vec{v_3} = \left(\frac{1}{\sqrt{2}} , 0, -\frac{1}{\sqrt{2}}\right)$, the co-ordinates of $\vec{u}$ relative to the basis S are given by $< \vec{u}, \vec{v_1} >$, $< \vec{u}, \vec{v_2} >$ and $< \vec{u}, \vec{v_3} >$.

We have $< \vec{u}, \vec{v_1} > = 2, < \vec{u}, \vec{v_2} > = \frac{4}{\sqrt{2}} = 2\sqrt{2}$, and $< \vec{u}, \vec{v_3} > = -\sqrt{2}$.

and $(\vec{u})_S = (< \vec{u}, \vec{v_1} >, < \vec{u}, \vec{v_2} >, < \vec{u}, \vec{v_3} >)$

$= (2, 2\sqrt{2}, -\sqrt{2})$.

Theorem 6 : If S is an orthonormal basis for an n-dimensional inner product space, and if $(\vec{u})_S = (u_1, u_2, \ldots, u_n)$ and $(\vec{v})_S = (v_1, v_2, \ldots, v_n)$, then

(a) $\| \vec{u} \| = \sqrt{u_1^2 + u_2^2 + \ldots\ldots + u_n^2}$

(b) $d(\vec{u}, \vec{v}) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$

(c) $< \vec{u}, \vec{v} > = u_1v_1 + u_2v_2 + \ldots + u_nv_n$.

Proof : Since $(\vec{u})_S = (u_1, u_2, \ldots u_n)$ and $(\vec{v})_S = (v_1, v_2, \ldots, v_n)$ if $S = \{\vec{w_1}, \vec{w_2}, \ldots, \vec{w_n}\}$ then by definition of co-ordinate vector,

$$\vec{u} = u_1\vec{w}_1 + u_2\vec{w}_2 + \ldots + u_n\vec{w}_n \text{ and}$$

$$\vec{v} = v_1\vec{w}_1 + v_2\vec{w}_2 + \ldots + v_n\vec{w}_n$$

(a) $\| \vec{u} \| = < \vec{u}, \vec{u} >^{1/2}$

$= [< u_1\vec{w_1} + u_2\vec{w_2} + \ldots + u_n\vec{w_n}, u_1\vec{w_1} + u_2\vec{w_2} + \ldots + u_n\vec{w_n} >]^{1/2}$

$= [< u_1\vec{w_1}, u_1\vec{w_1} + u_2\vec{w_2} + \ldots + u_n\vec{w_n} >$

$+ < u_2\vec{w_2}, u_1\vec{w_1} + \ldots + u_n\vec{w_n} > + \ldots + < u_n\vec{w_n}, u_1\vec{w_1} + \ldots + u_n\vec{w_n} >]^{1/2}$

$= [u_1^2 < \vec{w_1}, \vec{w_1} > + u_2^2 < \vec{w_2}, \vec{w_2} > + \ldots + u_n^2 < \vec{w_n}, \vec{w_n} >]^{1/2}$

$(\because < \vec{w_i}, \vec{w_j} > = 0, \text{ if } i \neq j)$

and $< \vec{w_i}, \vec{w_i} > = \| \vec{w_i} \|^2 = 1$, for each $i = 1, 2, \ldots n$

$\therefore \quad \| \vec{u} \| = [u_1^2 + u_2^2 + \ldots + u_n^2]^{1/2}$

(b) $d(\vec{u}, \vec{v}) = \| u - v \|$

$= [(u_1 - v_1) w_1 + (u_2 - v_2) w_2 + \ldots + (u_n - v_n) w_n]$

$= [(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots\ldots + (u_n - v_n)^2]^{1/2}$

(c) $< \vec{u}, \vec{v} > = < u_1w_1 + u_2w_2 + \ldots + u_nw_n, v_1w_1 + v_2w_2 + \ldots + v_nw_n >$

$= u_1v_1 + u_2v_2 + \ldots + u_nv_n$

Since $< \vec{w_i}, \vec{w_j} > = 0, \text{ if } i \neq j$ and $< \vec{w_i}, \vec{w_i} > = 1$.

Theorem 7 : Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of non-zero vectors in an inner product space. If all pairs of distinct vectors of S are orthogonal, then S is linearly independent.

Proof : Assume that

$$k_1\vec{v_1} + k_2\vec{v_2} + ... + k_n\vec{v_n} = 0 \qquad ... (1)$$

For each $\vec{v_i}$, we have from (1),

$$< k_1\vec{v_1} + k_2\vec{v_2} + ... + k_n\vec{v_n}, \vec{v_i} > = < 0, \vec{v_i} > = 0, \quad i = 1, 2, \ldots, n.$$

or equivalently,

$$k_1 < \vec{v_1}, \vec{v_i} > + k_2 < \vec{v_2}, \vec{v_i} > + ... \; k_n < \vec{v_n}, \vec{v_i} > = 0$$

From the orthogonality of S, it follows that $< v_j, v_i > = 0$, when $j \neq i$, so that this equation reduces to

$$k_i < \vec{v_i}, \vec{v_i} > = 0$$

Since $\vec{v_i} \neq 0$, $i = 1, 2, \ldots, n$, $< \vec{v_i}, \vec{v_i} > \neq 0$, so we must have $k_i = 0$, $i = 1, 2, \ldots, n$. This shows that S is linearly independent.

SOLVED EXAMPLES

Example 6.37 : Let $S = \{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ be set of vectors in $\mathbb{C}^3$, where $\vec{u_1} = (1, 0, i)$, $\vec{u_2} = (i, 0, 1)$ and $\vec{u_3} = (0, 1 + i, 0)$. Show that S is orthogonal set. Is S orthonormal ?

Solution : We have,

$$< \vec{u_1}, \vec{u_2} > = 1 \cdot \bar{i} + 0 \cdot \bar{0} + i \cdot \bar{1} = 1 \cdot (-i) + 0 + i \cdot 1 = 0$$

$$< \vec{u_1}, \vec{u_3} > = 0, \quad < \vec{u_2}, \vec{u_3} > = 0$$

$\therefore$ S is orthogonal set.

$$\|\vec{u_1}\| = (|1|^2 + |0|^2 + |i|^2)^{1/2} = (1 + 1)^{1/2} = \sqrt{2}$$

Similarly, $\|\vec{u_2}\| = \sqrt{2}$, $\|\vec{u_3}\| = \sqrt{2}$.

Since the vector in S are not unit vectors, so S is not orthonormal.

Orthogonal Projection onto a Line

Definition : Let V be an inner product space. Let $\vec{v}$ be a unit vector in V and $\vec{u} \in V$ is arbitary vector. Then the **projection** $P_{\vec{v}}(\vec{u})$ of $\vec{u}$ onto the line (one dimensional subspace $L(\vec{v}) = \{\alpha \vec{v} \mid \alpha \in \mathbb{R}\}$, is defined as $P_{\vec{v}}(\vec{u}) = < \vec{v}, \vec{u} > \vec{v}$. (See the following Fig. 6.3)

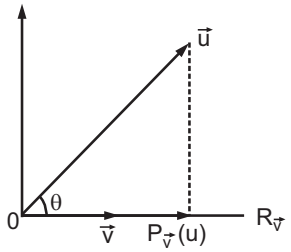


Fig. 6.3

From figure we see that $P_{\vec{v}}(\vec{u})$ is the point on $\mathbb{R}\vec{v}$ nearest to $\vec{u}$.

In this regards we prove the following theorem.

Theorem 8 : For a unit vector $\vec{v}$ and arbitrary vector $\vec{u}$ in V, let $P_{\vec{v}}(\vec{u}) = <\vec{u}, \vec{v}>\vec{v}$.

Then $d(P_{\vec{v}}(\vec{u}), \vec{u}) \leq d(\alpha\vec{v}, \vec{u})$ for any scalar α.

Proof : Observe from figure that $\vec{u} - P_{\vec{v}}(\vec{u})$ is orthogonal (perpendicular) to $\vec{v}$. We prove this algebraically.

$$\begin{aligned}
<\vec{u} - P_{\vec{v}}(\vec{u}), \vec{v}> &= <\vec{u}, \vec{v}> - <P_{\vec{v}}(\vec{u}), \vec{v}> \\
&= <\vec{u}, \vec{v}> - <<\vec{u}, \vec{v}>\vec{v}, \vec{v}> \\
&= <\vec{u}, \vec{v}> - <\vec{u}, \vec{v}><\vec{v}, \vec{v}> \\
&= <\vec{u}, \vec{v}> - <\vec{u}, \vec{v}> \quad (\because <\vec{v}, \vec{v}> = 1) \\
&= 0
\end{aligned}$$

Therefore, $\vec{u} - P_{\vec{v}}(\vec{u})$ is orthogonal to $\alpha\vec{v}$, for all $\alpha \mid \mathbb{R}$ and hence

$$(\vec{u} - P_{\vec{v}}(\vec{u}) \perp (P_{\vec{v}}(\vec{u}) - \vec{v}) \qquad \ldots (*)$$

Since $P_{\vec{v}}(\vec{u}) - \vec{v} \in \mathbb{R}_{\vec{v}}$

Now, for all $\alpha \in \mathbb{R}$, we have

$$\begin{aligned}
[d(\alpha\vec{v}, \vec{u})]^2 = \| \alpha\vec{v} - \vec{u} \|^2 &= \| \vec{u} - \alpha\vec{v} \|^2 \\
&= \| \vec{u} - P_{\vec{v}}(\vec{u}) + P_{\vec{v}}(\vec{u}) - \alpha\vec{v} \|^2 \\
&= \| \vec{u} - P_{\vec{v}}(\vec{u}) \|^2 + \| P_{\vec{v}}(\vec{u}) - \alpha\vec{v} \|^2 \\
&\qquad \text{(Using orthogonality (*))} \\
&\geq \| \vec{u} - P_{\vec{v}}(\vec{u}) \|^2 = [d(P_{\vec{v}}(\vec{u}), \vec{u})]^2
\end{aligned}$$

This gives, taking positive square roots :

$$d(\alpha \vec{v}, \vec{u}) \geq d(P_{\vec{v}}(\vec{u}), \vec{u})$$

Definition : Let $\vec{v}$ be a unit vector and $\vec{u}$ be an arbitrary vector is an inner product space V. We define the **orthogonal projection** of $\vec{u}$ along $\vec{v}$ by

$$P_{\vec{v}}(\vec{u}) = \langle \vec{u}, \vec{v} \rangle \vec{v}$$

If $\vec{v}$ is any non-zero vector, then we define

$$P_{\vec{v}}(\vec{u}) = \left\langle \vec{u}, \frac{\vec{v}}{\|\vec{v}\|} \right\rangle \frac{\vec{v}}{\|\vec{v}\|} = \langle \vec{u}, \vec{v} \rangle \frac{\vec{v}}{\|\vec{v}\|^2} = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \cdot \vec{v}$$

SOLVED EXAMPLES

Example 6.38 : Let $\vec{u} = (3, 1, -7)$ and $\vec{v} = (1, 0, 5)$. Find the orthogonal projection of $\vec{u}$ along $\vec{v}$.

Solution : We normalize $\vec{v}$, so that $\|\vec{v}\| = \sqrt{26}$.

$$\therefore \quad \vec{v'} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\sqrt{26}} \cdot \vec{v} = \left(\frac{1}{\sqrt{26}}, 0, \frac{5}{\sqrt{26}}\right) \text{ so that}$$

Then $$P_{\vec{v'}}(\vec{u}) = \langle \vec{u}, \vec{v'} \rangle \vec{v'}$$

$$\langle \vec{u}, \vec{v'} \rangle = \frac{3}{\sqrt{26}} + 0 + \frac{-35}{\sqrt{26}} = \frac{-32}{\sqrt{26}}$$

$$\therefore \quad P_{\vec{v'}}(\vec{u}) = \frac{-32}{\sqrt{26}} \vec{v'}$$

$$= -\frac{32}{\sqrt{26}} \left(\frac{1}{\sqrt{26}}, 0, \frac{5}{\sqrt{26}}\right)$$

$$= \left(\frac{-32}{26}, 0, \frac{-160}{26}\right)$$

$$= \left(\frac{-16}{13}, 0, \frac{-55}{13}\right)$$

The proofs of the following theorems are beyond the scope of this text, therefore we are going to state them only and will be used for application.

Theorem 9 : Let W be a finite dimensional subspace of an inner product space V.

(a) If $\{\vec{v_1}, \vec{v_2}, ..., \vec{v_r}\}$ is an orthonormal basis for W, and $\vec{u}$ is any vector in V, then

$$\text{Proj}_W \vec{u} = \langle \vec{u}, \vec{v_1} \rangle \vec{v_1} + \langle \vec{u}, \vec{v_2} \rangle \vec{v_2} + ... + \langle \vec{u}, \vec{v_r} \rangle \vec{v_r}.$$

(b) If $\{\vec{v_1}, \vec{v_2}, ..., \vec{v_r}\}$ is an orthogonal basis for W, and $\vec{u}$ is any vector in V, then

$$\text{Proj}_{W\vec{u}} = \frac{<\vec{u}, \vec{v_1}>}{\|\vec{v_1}\|^2}\vec{v_1} + \frac{<\vec{u}, \vec{v_2}>}{\|\vec{v_2}\|^2}\vec{v_2} + \ldots + \frac{<\vec{u}, \vec{v_r}>}{\|\vec{v_r}\|^2}\vec{v_r}$$

Definition : Let V be an inner product space and W be a finite dimensional subspace of V.

(a) Let $\{v_1, v_2, \ldots, v_r\}$ be *orthonormal* basis for W. Then for any $\vec{u}|V$, the *projection of* $\vec{u}$ *on W*, denoted by $\text{proj}_W(\vec{u})$ and is defined by the equation

$$\text{proj}_W \vec{u} = <\vec{u}, \vec{v_1}>\vec{v_1} + \ldots + <\vec{u}, \vec{v_r}>\vec{v_r}.$$

(b) Let $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_r}\}$ be an *orthogonal* basis for W. Then for each $\vec{u} \in V$, the *orthogonal projection* of $\vec{u}$ on W, denoted by $\text{proj}_{W\vec{u}}$, is defined by

$$\text{proj}_{W\vec{u}} = \frac{<\vec{u}, \vec{v_1}>}{\|\vec{v_1}\|^2}\vec{v_1} + \frac{<\vec{u}, \vec{v_2}>}{\|\vec{v_2}\|^2}\vec{v_2} + \ldots + \frac{<\vec{u}, \vec{v_r}>}{\|\vec{v_r}\|^2}\vec{v_r}$$

Definition : Let V be an inner product space and W be a finite dimensional subspace of V. For $\vec{u} \in V$, the orthogonal components of $\vec{u}$ to W is defined to be a vector $\vec{u} - \text{proj}_W \vec{u}$.

SOLVED EXAMPLES

Example 6.39 : Let $\mathbb{R}^3$ have the Euclidean inner product, and W be the subspace spanned by the orthogonal vectors $\vec{v_1} = (1, 0, 1)$ and $\vec{v_2} = (-1, 0, 1)$. If $\vec{u} = (1, 2, 3)$, find (a) orthogonal projection of $\vec{u}$ on W and (b) component of $\vec{u}$ orthogonal to W.

Solution : (a) From (b) of theorem 9, orthogonal projection of $\vec{u} = (1, 2, 3)$ on W is

$$\text{proj}_W(\vec{u}) = \frac{<\vec{u}, \vec{v_1}>}{\|\vec{v_1}\|^2}\vec{v_1} + \frac{<\vec{u}, \vec{v_2}>}{\|\vec{v_2}\|^2}\vec{v_2}$$

$$= \frac{1\times 1 + 2\times 0 + 3\times 1}{(1^2 + 0^2 + 1^2)}(1, 0, 1) + \frac{1\times(-1) + 2\times 0 + 3\times 1}{((-1)^2 + 0^2 + 1^2)}(-1, 0, 1)$$

$$= \frac{4}{2}(1, 0, 1) + \frac{2}{2}(-1, 0, 1) = (2, 0, 2) + (-1, 0, 1) = (1, 0, 3)$$

(b) Component of $\vec{u}$ orthogonal to W is

$$\vec{u} - \text{proj}_W(\vec{u}) = (1, 2, 3) - (1, 0, 3) = (0, 2, 0)$$

Remark : Observe that $\vec{u} - \text{proj}_W(\vec{u})$ is orthogonal to both $\vec{v_1}$ and $\vec{v_2}$, so that this vector is orthogonal to each vector in W spanned by $\vec{v_1}$ and $\vec{v_2}$.

Example 6.40 : Let $\vec{v_1} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ and $\vec{v_2} = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ be the orthonormal set of vectors in $\mathbb{R}^3$, with Euclidean inner product. If $\vec{u} = (1, 2, 3)$, compute.

(a) Orthogonal projection of $\vec{u}$ on w.

and (b) Orthogonal component of $\vec{u}$ to W.

Solution : (a) By (a) of theorem (9), we have

$$\begin{aligned}\text{proj}_W \vec{u} &= \langle \vec{u}, \vec{v_1} \rangle \vec{v_1} + \langle \vec{u}, \vec{v_2} \rangle \vec{v_2} \\ &= \frac{4}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + \frac{2}{\sqrt{2}}\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) \\ &= (2, 0, 2) + (-1, 0, 1) = (1, 0, 3).\end{aligned}$$

(b) Orthogonal component of u to W

$$\vec{u} - \text{proj}_W(\vec{u}) = (1, 2, 3) - (1, 0, 3) = (0, 2, 0).$$

We have the following theorem, which states that, every finite dimensional inner product space has an orthonormal basis and gives the step-by-step method of finding the orthonormal basis from the given basis of the vectors, which is known as *Gram-Schmidt process* of finding orthonormal basis.

Theorem 10 : Every non-zero finite dimensional inner product space has an orthonormal basis.

Let us describe Gram-Schmidt process for transforming a basis $\{\vec{v_1}, \vec{v_2}, ..., \vec{v_n}\}$ of an inner product space V to the orthogonal basis $\{\vec{u_1}, \vec{u_2}, ..., \vec{u_n}\}$ for V.

Step 1 : Set $\vec{u_1} = \vec{v_1}$. **Step 2 :** Let W_1 be subspace spanned by $\vec{u_1}$.

Take $\vec{u_2} = \vec{v_2} - \text{proj}_{W_1} \vec{v_2} = \vec{v_2} - \dfrac{\langle \vec{v_2}, \vec{u_1} \rangle}{\|\vec{u_1}\|^2} \vec{u_1}$.

Step 3 : Let W_2 be the subspace spanned by $\vec{u_1}$ and $\vec{u_2}$.

Take $$\vec{u_3} = \vec{v_3} - \text{proj}_{W_2} \vec{v_3} = \vec{v_3} - \left[\frac{\langle \vec{v_3}, \vec{u_1} \rangle}{\|\vec{u_1}\|^2} \vec{u_1} + \frac{\langle \vec{v_3}, \vec{u_2} \rangle}{\|\vec{u_2}\|^2} \vec{u_2}\right]$$

$$= \vec{v_3} - \frac{\langle \vec{v_3}, \vec{u_1} \rangle}{\|\vec{u_1}\|^2} \vec{u_1} - \frac{\langle \vec{u_3}, \vec{u_2} \rangle}{\|\vec{u_2}\|^2} \vec{u_2}$$

Step 4 : Let W_3 be the subspace spanned by $\vec{u_1}$, $\vec{u_2}$ and $\vec{u_3}$. For $\vec{u_4}$, take

$$\vec{u_4} = \vec{v_4} - \text{proj}_{w_3} \vec{v_4}$$

$$= \vec{v_4} - \left[\frac{< \vec{v_4}, \vec{u_1} >}{\|\vec{u_1}\|^2} \vec{u_1} + \frac{< \vec{v_4}, \vec{u_2} >}{\|\vec{u_2}\|^2} \vec{u_2} + \frac{< \vec{v_4}, \vec{u_3} >}{\|\vec{u_3}\|^2} \vec{u_3} \right]$$

Continuing in this way, we will obtain after n-steps, an orthogonal set of vectors $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$. Since V is n-dimensional and every orthogonal set is linearly independent, the set $\{\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}\}$ will be an orthogonal basis for V. Further normalising each of the vectors $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}$, the resulting set of vectors will be the orthonormal basis for V.

The step-by-step formation for converting an arbitrary basis into an orthogonal basis is called the *Gram-Schmidt process.*

Definition : Let V be a inner product space and S be a subset of V. The set

$$S^{\perp} = \{ \vec{u} \in V \mid < \vec{u}, \vec{w} > = 0, \text{ for all } \vec{w} \in S \}$$

is called S *perpendicular*. That is, in other words, $S^{\perp}$ is the set of all vectors in V which are perpendicular to each vector in S.

Theorem 11 : Let V be a finite dimensional inner product space and S be a subset of V. Then $S^{\perp}$ is a subspace of V.

Proof : Let $\vec{u}$, $\vec{v}$ be any two vectors in $S^{\perp}$ and α, β the scalars. Then, since $\vec{u}, \vec{v} \in S^{\perp}$, by definition $< \vec{u}, \vec{w} > = 0 = < \vec{v}, \vec{w} >$ for all $\vec{w} \in S$.

Now using properties of inner product,

$$< \alpha \vec{u} + \beta \vec{v}, \vec{w} > = < \alpha \vec{u}, \vec{w} > + < \beta \vec{v}, \vec{w} >$$

$$= \alpha < \vec{u}, \vec{w} > + \beta < \vec{v}, \vec{w} >$$

$$= \alpha \cdot 0 + \beta \cdot 0, \text{ for all } \vec{w} | S = 0$$

Thus, $< \alpha \vec{u} + \beta \vec{v}, \vec{w} > = 0$ for all $\vec{w} \in S$, hence $\alpha \vec{u} + \beta \vec{v} \in S^{\perp}$. Therefore, $S^{\perp}$ is a subspace of V.

Definition : Let V be a inner product space and W be a subspace of V. The W perpendicular $W^{\perp}$ is called *orthogonal complement of W.*

Note :

1. Clearly $W^{\perp}$ is a subspace by above theorem.
2. $W \cap W^{\perp} = \{\vec{0}\}$. For if $\vec{w} \in W \leftrightarrow W^{\perp}$, then $\vec{w} \in W$ and $\vec{w} \in W^{\perp}$, so $\vec{w}$ is orthogonal to each vector in W, in particular to itself. That is $< \vec{w}, \vec{w} > = 0$, but this holds only if $\vec{w} = \vec{0}$.

SOLVED EXAMPLES

Example 6.41 : Show that the set $\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right), \left(-\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)$ is orthogonal with respect to the Euclidean inner product. Convert it to an orthonormal set by normalising the vectors.

Solution : Set $\vec{v_1} = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$, $\vec{v_2} = \left(-\frac{1}{2}, \frac{1}{2}, 0\right)$, and $\vec{v_3} = \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)$. Then

$$< \vec{v_1}, \vec{v_2} > = \left(\frac{1}{5}\right)\left(\frac{-1}{2}\right) + \left(\frac{1}{5}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{5}\right)(0)$$

$$= -\frac{1}{10} + \frac{1}{10} + 0 = 0$$

$$< \vec{v_2}, \vec{v_3} > = \left(-\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + (0)\left(\frac{2}{3}\right)$$

$$= -\frac{1}{6} + \frac{1}{6} + 0 = 0$$

and $$< \vec{v_1}, \vec{v_3} > = \left(\frac{1}{5}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{5}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{5}\right)\left(-\frac{2}{3}\right)$$

$$= \frac{1}{15} + \frac{1}{15} - \frac{2}{15} = 0.$$

Therefore $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$ are orthogonal vectors.

Now,

$$\| \vec{v_1} \| = \sqrt{\left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2} = \sqrt{\frac{3}{25}} = \frac{1}{5}\sqrt{3}$$

$$\| \vec{v_2} \| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + 0^2} = \sqrt{\frac{2}{4}} = \frac{1}{\sqrt{2}}.$$

and $$\| \vec{v_3} \| = \sqrt{\left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2}$$

$$= \sqrt{\frac{6}{9}} = \sqrt{\frac{2}{3}}$$

The orthonormal set is obtained by converting $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$ to $\frac{1}{\|\vec{v_1}\|}\vec{v_1}$, $\frac{1}{\|\vec{v_2}\|}\vec{v_2}$ and $\frac{1}{\|\vec{v_3}\|}\vec{v_3}$ respectively. We have the orthonormal set

$$\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)\right\}.$$

Example 6.42 : Show that the vectors $\overrightarrow{v_1} = \left(-\frac{3}{5}, \frac{4}{5}, 0\right)$, $\overrightarrow{v_2} = \left(\frac{4}{5}, \frac{3}{5}, 0\right)$, $\overrightarrow{v_3} = (0, 0, 1)$ form an orthonormal basis for $\mathbb{R}^3$ with Euclidean inner product. Find the co-ordinates of (1, –1, 2) relative to this basis.

Solution : We have

$$< \overrightarrow{v_1}, \overrightarrow{v_2} > = \left(-\frac{3}{5}\right)\left(\frac{4}{5}\right) + \left(\frac{4}{5}\right)\left(\frac{3}{5}\right) + 0 = 0$$

$$< \overrightarrow{v_2}, \overrightarrow{v_3} > = \left(\frac{4}{5}\right)(0) + \left(\frac{3}{5}\right)(0) + 0 \times 1 = 0$$

and $$< \overrightarrow{v_1}, \overrightarrow{v_3} > = \left(-\frac{3}{5}\right)(0) + \left(\frac{4}{5}\right)(0) + (0)(1) = 0$$

Also, $$\| \overrightarrow{v_1} \| = \sqrt{\left(-\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 + (0)^2} = \sqrt{\frac{9}{25} + \frac{16}{25}} = 1$$

$$\| \overrightarrow{v_2} \| = \sqrt{\left(\frac{4}{5}\right)^2 + \left(\frac{3}{5}\right)^2 + (0)^2} = 1$$

and $$\| \overrightarrow{v_3} \| = \sqrt{0^2 + 0^2 + 1^2} = 1.$$

Thus $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ and $\overrightarrow{v_3}$ are orthogonal and each is of norm 1, hence is an orthonormal basis for $\mathbb{R}^3$, since orthogonal set is linearly independent and $\mathbb{R}^3$ has dimension 3.

To find the co-ordinates of $\overrightarrow{u} = (1, -1, 2)$ relative to this basis, we know from the remark below the theorem (7), that the co-ordinates of $\overrightarrow{u}$ are given by $< \overrightarrow{u}, \overrightarrow{v_1} >$, $< \overrightarrow{u}, \overrightarrow{v_2} >$ and $< \overrightarrow{u}, \overrightarrow{v_3} >$. Then we have,

$$< \overrightarrow{u}, \overrightarrow{v_1} > = (1)\left(-\frac{3}{5}\right) + (-1)\left(\frac{4}{5}\right) + (2)(0) = -\frac{9}{5}$$

$$< \overrightarrow{u}, \overrightarrow{v_2} > = (1)\left(\frac{4}{5}\right) + (-1)\left(\frac{3}{5}\right) + (2)(0) = \frac{1}{5}$$

and $$< \overrightarrow{u}, \overrightarrow{v_3} > = (1)(0) + (-1)(0) + (2)(1) = 2.$$

Thus, $-\frac{9}{5}, \frac{1}{5}, 2$ are co-ordinates of $\overrightarrow{u}$ relative to this basis.

Example 6.43 : Let $S = \{\overrightarrow{w_1}, \overrightarrow{w_2}\}$ be orthonormal basis for $\mathbb{R}^2$ with Euclidean inner product, where $\overrightarrow{w_1} = \left(\frac{3}{5}, -\frac{4}{5}\right)$ and $\overrightarrow{w_2} = \left(\frac{4}{5}, \frac{3}{5}\right)$. Find the vector u having co-ordinate vector $(\overrightarrow{u})_S = (1, 1)$. Verify part (a) of theorem 8.

Solution : Let $\overrightarrow{u} = (a, b)$ be the vector such that $(\overrightarrow{u})_S = (1, 1)$.

Now, by remark of theorem 7, we have,

$$(\vec{u})_s = (1, 1) = (< \vec{u}, \vec{w_1} >, < \vec{u}, \vec{w_2} >)$$
$$= \left(\frac{3}{5}a - \frac{4}{5}b, \frac{4}{5}a + \frac{3}{5}b\right)$$

Equating the corresponding components,

$$\frac{3}{5}a - \frac{4}{5}b = 1$$

and $$\frac{4}{5}a + \frac{3}{5}b = 1.$$

Solving for a and b, we obtain, $a = \frac{7}{5}$ and $b = -\frac{1}{5}$.

Thus $$\vec{u} = \left(\frac{7}{5}, -\frac{1}{5}\right).$$

Now, by theorem 8, if $(\vec{u})_s = (\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n})$, then

$\| \vec{u} \| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$. So here $(\vec{u})_s = (1, 1)$.

$\therefore$ $\| \vec{u} \| = \sqrt{1^2 + 1^2} = \sqrt{2}$. Also as $\vec{u} = \left(\frac{7}{5}, \frac{1}{5}\right)$

we have $$\| \vec{u} \| = \sqrt{\left(\frac{7}{5}\right)^2 + \left(-\frac{1}{5}\right)^2} = \sqrt{2}.$$

Thus the theorem is verified.

Example 6.44 : Let $\mathbb{R}^2$ and $\mathbb{R}^3$ have Euclidean inner product. Use Gram-Schmidt process to transform

(a) the basis $\vec{u_1} = (1, -3)$, and $\vec{u_2} = (2, 2)$ to orthonormal basis of $\mathbb{R}^2$ and

(b) the basis $\vec{u_1} = (1, 1, 1)$, $\vec{u_2} = (-1, 1, 0)$ and $\vec{u_3} = (1, 2, 1)$ to orthonormal basis of $\mathbb{R}^3$.

Solution : (a) Set $\vec{v_1} = \vec{u_1} = (1, -3)$ and $W_1 = L\{(\vec{v_1})\}$.

Then $$\vec{v_2} = \vec{u_2} - \text{proj}_{W_1} \vec{u_2} = \vec{u_2} - \frac{< \vec{v_2}, \vec{v_1'} >}{\| \vec{v_1'} \|^2} \vec{v_1'}$$

Here $\vec{u_2} = (2, 2)$, $\| \vec{v_1} \|^2 = 10$, $< \vec{u_2}, \vec{v_1} > = -4$

$\therefore$ $$\vec{v_2} = (2, 2) - \left(-\frac{4}{10}\right)(1, -3)$$
$$= (2, 2) + \left(\frac{4}{10}, -\frac{12}{10}\right) = \left(\frac{12}{5}, \frac{4}{5}\right).$$

Note $\vec{v}_1$ and $\vec{v}_2$ are orthogonal to each other.

$$\|\vec{v}_2\| = \sqrt{\left(\frac{12}{5}\right)^2 + \left(\frac{4}{5}\right)^2} = \sqrt{\frac{144+16}{25}} = 4\sqrt{\frac{2}{5}}\ .$$

Normalising $\vec{v}_1$ and $\vec{v}_2$, we get

$$\vec{v}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{10}}(1, -3) = \left(\frac{1}{\sqrt{10}}, -\frac{3}{\sqrt{10}}\right)$$

and

$$\vec{v}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{1}{4\sqrt{\frac{2}{5}}}\left(\frac{12}{5}, \frac{4}{5}\right) = \left(\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right).$$

So that the set $\{\vec{v}_1, \vec{v}_2\}$ is an orthonormal basis for $\mathbb{R}^2$.

(b) **Step 1 :** Take $\vec{v}_1 = \vec{u}_1 = (1, 1, 1)$

Step 2 : Let $W_1 = L\{\vec{u}_1\}$ and take

$$\vec{v}_2 = \vec{u}_2 - \text{proj}_{W_1}\vec{u}_2 = \vec{u}_2 - \frac{\langle \vec{v}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2}\vec{v}_1$$

We have $\langle \vec{u}_2, \vec{v}_1 \rangle = (1)(-1) + (1)(1) + (1)(0) = 0$

$$\|\vec{v}_1\|^2 = 3,\ \vec{u}_2 = (-1, 1, 0)$$

$$\vec{v}_2 = (-1, 1, 0) - \frac{0}{3}\vec{v}_1 = (-1, 1, 0).$$

Step 3 : Let $W_2 = L\{\vec{v}_1, \vec{v}_2\}$. Take

$$\vec{v}_3 = \vec{u}_3 - \text{proj}_{W_2}\vec{u}_3$$

that is,

$$\vec{v}_3 = \vec{u}_3 - \frac{\langle \vec{u}_3, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2}\vec{v}_1 - \frac{\langle \vec{u}_3, \vec{v}_2 \rangle}{\|\vec{v}_2\|^2}\vec{v}_2$$

or

$$\langle \vec{u}_3, \vec{v}_1 \rangle = (1)(1) + (2)(1) + (1)(0) = 1.$$

$\| \vec{v}_2 \|^1 = \sqrt{2}$. Substituting in (*), we get

$$\vec{v}_3 = (1, 2, 1) - \left(\frac{4}{3}\right)(1, 1, 1) - \frac{1}{2}(-1, 1, 0)$$

$$= \left(1 - \frac{4}{3} + \frac{1}{2}, 2 - \frac{4}{3} - \frac{1}{2}, 1 - \frac{4}{3}\right)$$

$$= \left(\frac{1}{6}, \frac{1}{6}, -\frac{1}{3}\right).$$

Note that the vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are pairwise orthogonal and hence are linearly independent, therefore form a basis for $\mathbb{R}^3$. Now, normalising $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$, we get

$$\vec{w}_1 = \frac{\vec{v}_1}{\| \vec{v}_1 \|} = \frac{1}{\sqrt{3}}(1, 1, 1) = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$

$$\vec{w}_2 = \frac{\vec{v}_2}{\| \vec{v}_2 \|} = \frac{1}{\sqrt{2}}(-1, 1, 0) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$$

and

$$\vec{w}_3 = \frac{\vec{v}_3}{\| \vec{v}_3 \|} = \sqrt{6}\left(\frac{1}{6}, \frac{1}{6}, -\frac{1}{3}\right)$$

$$= \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\sqrt{\frac{2}{3}}\right)$$

Thus $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ is an orthonormal basis for $\mathbb{R}^3$.

Example 6.45 : Resolve $\vec{u}$ into two perpendicular components along $\vec{v}$, where $\vec{u} = (1, i + 1, 0)$ and $\vec{v} = (i - 1, i, 1)$.

Solution : We know that the projection of $\vec{u}$ along $\vec{v}$; $\text{proj}_{\vec{v}}\, \vec{u}$ is given by

$$\text{proj}_{\vec{v}}\, \vec{u} = \frac{< \vec{u}, \vec{v} >}{< \vec{v}, \vec{v} >} \vec{v} \qquad \ldots (*)$$

Now, $$\begin{aligned} \langle \vec{u}, \vec{v} \rangle &= 1 \cdot \overline{(i-1)} + (i+1)\,\bar{i} + 0 \cdot \bar{1} \\ &= 1\,(-i-1) + (i+1)\,(-i) + 0 \\ &= -i - 1 + 1 - i + 0 \\ &= -2i \end{aligned}$$

$$\begin{aligned} \langle \vec{v}, \vec{v} \rangle &= (|i-1|^2 + |i|^2 + |1|^2) \\ &= (2 + 1 + 1) \\ &= 4 \end{aligned}$$

Using these values in (*), we get

Using these values in (*), we obtain

$$\text{proj}_{\vec{v}}\,\vec{u} = \frac{-2i}{4}(i-1, i, 1) = \frac{1}{2}(1+i, 1, -i)$$

And the orthogonal projection of $\vec{u}$ to $\vec{v}$ is

$$\begin{aligned} \vec{u} - \text{proj}_{\vec{v}}\,\vec{u} &= (1, i+1, 0) - \frac{1}{2}(1+i, 1, -i) \\ &= \frac{1}{2}(1-i, 1+2i, i) \end{aligned}$$

Example 6.46 : Resolve $\vec{u} = (1 + i, 2 + i, -5, -6)$ into two perpendicular components along the vector $\vec{v} = (1, 2 + i, i, 1 - i)$.

Solution : First component is $\text{proj}_{\vec{v}}\,\vec{u}$, which is given by

$$\begin{aligned} \text{proj}_{\vec{v}}\,\vec{u} &= \langle \vec{u}, \vec{v} \rangle \vec{v} \\ &= [(1+i)\,1 + (2+i)\,(2-i) - 5\,(-i) + (-6)\,(1+i)]\,\vec{v} \\ &= [1 + i + 5 + 5i - 6 - 6i]\,\vec{v} \\ &= \vec{0} \end{aligned}$$

Therefore, this shows that $\vec{u}$ is itself perpendicular to $\vec{v}$, hence its two perpendicular components are : one $\vec{v}$ and the other is $\vec{u}$ itself.

Example 6.47 : Use Gram-Schmidt process to convert the basis $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ to an orthonormal basis $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$, where $\vec{v_1} = (1, 0, 0)$, $\vec{v_2} = (1, 1, 0)$, $\vec{v_3} = (1, 1, 1)$.

Solution : We first transform these vectors to $\vec{u'_1}, \vec{u'_2}, \vec{u'_3}$ which are mutually orthogonal. For this Gram-Schmidt process says that :

Take $\vec{u'_1} = \vec{v_1} = (1, 0, 0)$

Next $\vec{u'_2} = v_2 - \text{proj}_{\vec{u'_1}} \vec{v_2}$

$$= (1, 1, 0) - \frac{< \vec{v'_2}, \vec{u'_1} >}{< \vec{u'_1}, \vec{u'_1} >} \vec{u'_1}$$

$$= (1, 1, 0) - \frac{1}{1} \vec{u'_1} = (0, 1, 0)$$

since $< \vec{v'_2}, \vec{v'_1} > 1, \quad < \vec{u'_1}, \vec{u'_1} > = 1$

Further, $\vec{u'_3} = \vec{v_3} - \text{proj}_{\vec{u'_1}} \vec{v_3} - \text{proj}_{\vec{u'_2}} \vec{v_3}$

$$= \vec{v_3} - \frac{< \vec{v_3}, \vec{u'_1} >}{< \vec{u'_1}, \vec{u'_1} >} \vec{u'_1} - \frac{< \vec{v_3}, \vec{u'_2} >}{< \vec{u'_2}, \vec{u'_2} >} \vec{u'_2} \qquad \ldots (*)$$

Now, $< \vec{v_3}, \vec{u'_1} > = 1, \ < \vec{u'_1}, \ \vec{u'_1} > = 1$;

$< \vec{v_3}, \vec{u'_2} > = 1, < \vec{u'_2}, \ \vec{u'_2} > = 1$, hence using these all in (*), we obtain

$$\vec{u'_3} = (1, 1, 1) - \vec{u'_1} - \vec{u'_2}$$
$$= (1, 1, 1) - (1, 0, 0) - (0, 1, 1)$$
$$= (0, 0, 1)$$

Observe here that all $\vec{u'_1}, \vec{u'_2}$ and $\vec{u'_3}$ are mutually orthogonal and unit vectors, hence we need normalize $\vec{u'_1}, \vec{u'_2}$ and $\vec{u'_3}$ as the last step. Thus, the set $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ is orthonomal basis for $\mathbb{R}^3$, where $\vec{u_1} = (1, 0, 0), \vec{u_2} = (0, 1, 0), \vec{u_3} = (0, 0, 1)$.

Example 6.48 : Let $\vec{v_1} = (i, 0, 1), \vec{v_2} = (i - 1, i, 1)$ and $\vec{v_3} = (1, i + 1, 0)$. Show that $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ is a basis for $\mathbb{C}^3$. Further transform this basis to an orthonormal basis $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ by using Gram-Schmidt process.

Solution : As usual the first step is to take

$$\vec{u'_1} = \vec{v_1} = (i, 0, 1)$$

Second step is to find $\vec{u'_2}$, where,

$$\vec{u'_2} = \vec{v_2} - \text{proj}_{\vec{u'_1}} \vec{v_2}$$

$$= \vec{v_2} - \frac{< \vec{v_2}, \vec{u'_1} >}{< \vec{u'_1}, \vec{u'_1} >} \vec{u'_1} \qquad (*)$$

We have, $< \vec{v_2}, \vec{u'_1} > = (i - 1)\,\bar{i} + i \times 0 + 1 \cdot \bar{1}$

$$= + 1 + i + 0 + 1 \qquad (\because \bar{i} = - i)$$
$$= 2 + i$$

and $< \vec{u'_1}, \vec{u'_1} > = 2$

Using these values in (*) we obtain

$$\therefore \quad \vec{u'_2} = (i - 1, i, 1) - \frac{2 + i}{2}(i, 0, 1)$$

$$= (i - 1, i, 1) - \frac{1}{2}(2i - 1, 0, 2 + i) = \left(- \frac{1}{2}, i, - \frac{i}{2} \right)$$

We can take $\vec{u'_2} = (1, - 2i, i)$, because if a vector $\vec{u}$ is perpendicular to $\vec{v}$, then any scalar multiple of $\vec{u}$ is also perpendicular to $\vec{v}$.

Thus, $\vec{u'_2} = (1, - 2i, i)$

Third step is to find $\vec{u'_3}$, we have

$$\vec{u'_3} = \vec{v_3} - \text{proj}_{\vec{u'_1}} \vec{v_3} - \text{proj}_{\vec{u'_2}} \vec{v}$$

$$= \vec{v_3} - \frac{< \vec{v_3}, \vec{u'_1} >}{< \vec{u'_1}, \vec{u'_1} >} \vec{u'_1} - \frac{< \vec{v_3}, \vec{u'_2} >}{< \vec{u'_2}, \vec{u'_2} >} \vec{u'_2} \qquad \ldots (**)$$

Now, $< \vec{v_3}, \vec{u'_1} > = 1\,(\bar{i}) + (i + 1)\,\bar{0} + 0 \cdot \bar{1} = - i$

$< \vec{u'_1}, \vec{u'_1} > = 2$

$< \vec{v_3}, \vec{u'_2} > = 1 \cdot 1 + (i + 1)\,(\overline{- 2i}) + 0 \cdot \bar{i} = - 1 + 2i$

and $< \vec{u'_2}, \vec{u'_2} > = |1|^2 + |- 2i|^2 + |i|^2 = 6$

Using these values in (**), we obtain

$$\vec{u'_3} = (1, i + 1, 0) - \frac{1}{2}(1, 0, - i) - \frac{1}{6}(- 1 + 2i, 4 + 2i - i - 2)$$

$$= (2 - i, 2i + 1, 2i + 1)$$

$$\frac{\langle \vec{v_3}, \vec{u'_1} \rangle}{\langle \vec{u'_1}, \vec{u'_1} \rangle} \vec{u'_1} = \frac{1}{2}(1, 0, -i)$$

where, $$\frac{\langle \vec{u_3}, \vec{u'_2} \rangle}{\langle \vec{u'_2}, \vec{u'_2} \rangle} \vec{u'_2} = \frac{1}{6}(-1 + 2i, 4 + 2i, -i - 2)$$

Thus, we obtained $\vec{u'_1}, \vec{u'_2}, \vec{u'_3}$ which are mutually orthogonal. Now we normalize each of these vectors. Since

$$\|\vec{u'_1}\| = \sqrt{2}, \ \|\vec{u'_2}\| = \sqrt{6} \text{ and } \|\vec{u'_3}\| = \sqrt{15}$$

Therefore, the orthonormal basis obtained is $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$, where,

$$\vec{u_1} = \frac{1}{\sqrt{2}}(i, 0, 1), \ \vec{u_2} = \frac{1}{\sqrt{6}} = (1, -2i, i)$$

and $$\vec{u_3} = \frac{1}{\sqrt{15}}(2 - i, 2i + 1, 2i + 1)$$

6.6 Rank and Nullity : Sylvester Inequality

We know that the rank of a matrix when premultiplied or post multiplied by a non-singular matrix is unaltered. The rank of the product of two or more matrices were studied by using the basic concept of elementary row or column transformations.

Here in this section we find the same results by using the concept of range space and null space of a matrix. The most important part of this section is the **Sylvester inequality**, which gives upper and lower bound for the rank of product of two matrices.

Definition : Let A be an $m \times n$ matrix and X be n-vector considered as column matrix. The set of all solutions of $AX = 0$ forms a vector space, which we called **solution space** of A or it is also defined as the **null space** of A and is denoted by N(A). The set of all vectors Y in $\mathbb{R}^m$ such that $AX = Y$, $Y \neq 0$, for some X in $\mathbb{R}^n$ is defined as the **range space** of A and is denoted by R(A). The dimension of null space N(A) is called **nullity of A** and is denoted by **null (A)**. The dimension of the range space R(A) is **called rank of A** and is denoted by rank (A).

Theorem 11 : Any vector Y in range space of A is linear combination of the columns of A.

Proof : Let A be $m \times n$ matrix, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ : & : & & : \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Let Y be any vector in R(A), so that by definition there is some $X = \begin{bmatrix} x_1 \\ x_2 \\ : \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$ such that $AX = Y$

or $\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ : & : & & : \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ : \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ : \\ y_m \end{bmatrix}$ say.

Then matrix multiplication on left side gives

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ : \quad : \quad : \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ : \\ y_n \end{bmatrix}$$

Again the left hand column matrix can be written as :

$$\begin{bmatrix} a_{11} \\ a_{21} \\ : \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ : \\ a_{m2} \end{bmatrix} x_2 + \dots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ : \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} y_1 \\ y_2 \\ : \\ y_m \end{bmatrix}$$

which shows that Y is a linear combination of column vectors of A.

Remark : As R(A) is the set of all linear combinations of columns of A, the range space of A is spanned by the columns vectors of A. Therefore, the number of linearly independent column vectors of A form a basis for range space of A.

SOLVED EXAMPLES

Example 6.49 : Find the null space and range space in each of : (a) A = [– 1, 2, 1], (b) $A = \begin{bmatrix} 0 & 2 & 2 \\ 1 & -1 & 1 \end{bmatrix}$.

Solution : (a) The null space of A is given by

$$[-1,\ 2,\ 1] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

This means N(A) is the set of those vectors in $\mathbb{R}^3$, which are orthogonal to (– 1, 2, 1), so it is clear that N(A) is plane passing through origin and orthogonal to the vector (– 1, 2, 1). Hence null (A) = 2. Also the range space of A is clearly the set of real numbers, hence rank A = 1. Thus rank (A) + null (A) = 3.

(b) The null space of A is given by

$$\begin{bmatrix} 0 & 2 & 2 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In this case the null space is the set of vectors which are orthogonal to both the vectors (0, 2, 2) and (1, − 1, 1) hence must be a line passing through origin and perpendicular to the plane containing the vectors (0, 2, 2) and (1, − 1, 1). Hence, null A = 1. The range space of A is the set of vectors in $\mathbb{R}^3$, which are linear combinations of column vectors of A, clearly there are **two** linearly independent column vectors of A, therefore rank A = 2. Again in this case

$$\text{rank (A)} + \text{null (A)} = 2 + 1 = 3.$$

Theorem 12 : If A is m × n matrix, then rank of A is r if and only if null (A) = n − r.

Proof : Let rank A = r. Then only solution of AX = 0 is of the form $X = \begin{bmatrix} u \\ w \end{bmatrix}$, where u is an r-vector and w is an arbitrary (n − r) – vector. For i = 1, 2, …, n − r, let e_i be standard basis vectors in $\mathbb{R}^{n-r}$ such that

$$w = \sum_{i=1}^{n-r} w_i e_i$$

where, w_i are components of w and let b_i be linear combination of e_i, i = 1, 2, …, r. Then

$$X = \begin{bmatrix} u \\ w \end{bmatrix} = \sum_{i=1}^{n-r} w_i \begin{bmatrix} b_i \\ e_i \end{bmatrix} = \sum_{i=1}^{n-r} w_i y_i, \text{ where } y_i = \begin{bmatrix} b_i \\ e_i \end{bmatrix}.$$

Then (n − r) number of vectors Y_i are linearly independent and solution of AX = 0 can always be represented in terms of (n − r) number of vectors. Therefore, the set of Y_i's, i = 1, 2, …, n − r is basis of N(A) and null (A) = n − r.

Conversely, suppose null (A) = n − r. Then there exist (n − r) linearly independent vectors X_i's in $\mathbb{R}^n$ such that $AX_i = 0$ for all i = r + 1, …, …, n. We denote by X_k's, k = 1, 2, …, r, the remaining independent vectors in $\mathbb{R}^n$, so that r X_i's and X_k's together form a basis of $\mathbb{R}^n$. So we have $AX_k = Y_k \neq 0$ for k = 1, 2, …, r, where Y_k's are in the range space R(A) of A.

Let $\alpha_1 Y_1 + \alpha_2 Y_2 + \ldots + \alpha_r Y_r = 0$

Then substituting for Y_k's, we get

$$A(\alpha_1 X_1 + \ldots + \alpha_r X_r) = 0 \text{ or } AX = 0$$

where, $X = \alpha X_1 + \ldots + \alpha X_2$ is not in N(A). Hence X must be equal to the zero. Since X_k's are linearly independent $\alpha_1 = \ldots = \alpha_r = 0$. Thus Y_k's are also linearly independent, so

$$r = \dim R(A) = \text{rank (A)}$$

Remark : In above proof some steps are taken for granted as, these steps require many results which are not covered in this text. The same thing can happen in following results in this section.

Corollary : Let A be an $m \times n$ matrix, then

$$\text{rank (A)} + \text{null (A)} = n$$

Proof : The proof follows clearly from above theorem.

Theorem 13 : If A and B are two matrices respectively of order $m \times n$ and $n \times p$, then

$$\text{rank (AB)} \leq \min (\text{rank A, rank B})$$

Proof : Let X be a vector in $\mathbb{R}^P$. Then BX = 0 implies (AB) X = 0, hence $N(B) \leq N(AB)$. Therefore, null (B) ≤ null (AB).

Also, we know p-rank (B) ≤ P − rank (AB), or

$$\text{rank (AB)} \leq \text{rank (B)} \quad \ldots \text{(i)}$$

Similarly, we can show that

$$\text{rank (B'A')} \leq \text{rank (A')},$$

hence $$\text{rank (AB)} \leq \text{rank (A)} \quad \ldots \text{(ii)}$$

Combining equation (i) and (ii), we get that

$$\text{rank (AB)} \leq \min (\text{rank (A), rank (B)})$$

Corollary : If P is is $m \times m$ and Q is $n \times n$ non-singular matrices, then

$$\text{rank (PA)} = \text{rank(A) and rank (AQ) = rank (A)},$$

where, A is $m \times n$ matrix.

Proof : By theorem (B), rank (PA) ≤ rank (A) and rank (A) = rank (P^{-1}PA) ≤ rank (PA), which proves that rank (PA) = rank A. Similarly we can prove rank (AQ) = rank (A).

Using above result it can be proved.

Corollary : If P and Q are non-singular matrices, then rank (A) = rank (PAQ), where A is $m \times n$ matrix and P, Q as in above corollary.

Theorem 14 : If A and B are matrices of order $m \times n$,

$$\text{then rank (A + B)} \leq \text{rank (A)} + \text{rank (B)}.$$

Proof : Let rank (A) = r and rank (B) = p. Let $X_1, X_2, \ldots, X_r$ and $Y_1, Y_2, \ldots, Y_p$ be the independent columns of A and B respectively. The columns $Z_1, Z_2, \ldots, Z_n$ of (A + B), which a vector space W, say can be expressed as linear combination of X_i's and Y_i's and therefore, for all k = 1, …, n, Z_k is in W implies that Z_k too is in vector space V spanned by X_i's and Y_i's.

So W is a subspace of V. Hence

$$\dim W \leq \dim V \leq r + p$$

because the number of independent vectors along X_i's and Y_i's is less than or equal to r + p. Hence, rank (A + B) ≤rank (A) + rank (B)

Theorem 15 : Let A and B be two matrices of order $m \times n$ and $n \times p$ respectively. Then the number of linearly independent vectors Y = BX satisfying AY = 0 is rank (B) − rank (AB).

Proof : Let null (B) = p − rank (B) = q and null (AB) = p − rank (AB) = k

So there exist q number of linearly independent vectors : $X_1, X_2, \ldots, X_q$ in $\mathbb{R}^p$ satisfying $BX = 0$ which is the basis of N(B). Now every vector satisfying $BX = 0$ also satisfies $(AB)\ X = 0$. So, N(B) is a subspace of N(AB) and hence $k \geq q$. From this it follows that in $\mathbb{R}^p$ there are $(k - q)$ number of independent vectors $X_{q+1}, \ldots, X_k$ such that $X_1, X_2, \ldots, X_k$ is a basis of N(AB) and $BX_i \neq 0$, $i = q = 1, \ldots, k$. We have to show that $BX_{q+1}, \ldots, BX_k$ are linearly independent. Suppose we have scalars $\alpha_1, \ldots, \alpha_q$ such that

$$\alpha_{q+1} X_{q+1} + \ldots + \alpha_k X_k = \alpha_1 X_1 + \ldots + \alpha_q X_q$$

Since $X_1, \ldots, X_q, X_{q+1}, \ldots, X_k$ are linearly independent, $\alpha_1, \ldots, \alpha_q, \alpha_{q+1}, \ldots, \alpha_k$ are all zero. As a consequence it follows that $BX_{q+1}, \ldots, BX_k$ are linearly independent. Hence, the number of independent vectors $Y = BX$ satisfying $AY = 0$ is

$$k - q = \text{rank } (B) - \text{rank } (AB);$$

that is, null (AB) = k + rank (B) − rank (AB).

Theorem 16 : (Sylvester Inequality) : If A and B are matrices of order $m \times n$ and $n \times p$ respectively, then

$$\text{rank } A + \text{rank } B - n \leq \text{rank } (AB) \leq \min (\text{rank } A, \text{rank } B)$$

Proof : We see that n − rank (A) = null (A); that is, the number of independent vectors Y satisfying $AY = 0$. Again, we have rank B − rank (AB) is the number of independent vectors $Y = BX$ satisfying $AY = 0$ and forms a subspace of N(A). Hence,

$$\text{rank } B - \text{rank } (AB) \leq n - \text{rank } A,$$

$$\text{rank } A + \text{rank } B - n \leq \text{rank } (AB)$$

We have already prove second inequality and so the proof of the theorem follows.

Theorem 17 : If A, B and C are matrices of order $m \times n$, $n \times p$ and $p \times k$ respectively, then

$$\text{rank } (AB) + \text{rank } (BC) \leq \text{rank } B + \text{rank } (ABC).$$

Proof : We note that the number of independent vectors BCX such that $(ABC)\ X = 0$ is [rank (BC) − rank (ABC)] and span a vector space V and the number of independent vectors BX satisfying $ABX = 0$ is [rank B − rank (AB)] and span a vector space W. Since V is a subspace of W.

$$\text{rank } (BC) - \text{rank } (ABC) \leq \text{rank } B - \text{rank } (AB).$$

Exercise 6.2

1. Let P_2 have the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)\, q(x)\, dx$. Show that $p(x) = x$ and $q(x) = x^2$ are orthogonal.

2. If P_2 has an inner product as in exercise 1, and if p (x) = x and q (x) = x^2, verify Pythagorus theorem.
3. Let $\mathbb{R}^4$ have Euclidean inner product, and let $\vec{u}$ = (–1, 1, 0, 2). Determine whether the vector $\vec{v}$ is orthogonal to the set of vectors W = $\{\vec{w_1}, \vec{w_2}, \vec{w_3}\}$, where $\vec{w_1}$ = (0, 0, 0, 0), $\vec{w_2}$ = (1, –1, 3, 0) and $\vec{w_3}$ = (4, 0, 9, 2).
4. Let $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^4$ have Euclidean inner product. In each part, find the cosine of the angle between $\vec{u}$ and $\vec{v}$.
   (a) $\vec{u}$ = (1, – 3), $\vec{v}$ = (2, 4)
   (b) $\vec{u}$ = (–1, 0), $\vec{v}$ = (3, 8)
   (c) $\vec{u}$ = (–1, 5, 2), $\vec{v}$ = (2, 4, –9)
   (d) $\vec{u}$ = (4, 1, 8), $\vec{v}$ = (1, 0, –3)
   (e) $\vec{u}$ = (1, 0, 1, 0), $\vec{v}$ = (–3, –3, –3, –3)
   (f) $\vec{u}$ = (2, 1, 7, –1), $\vec{v}$ = (4, 0, 0, 0).
5. Let P_2 have the inner product as in exercise 1. Find the cosine of the angle between p and q.
   (a) $p = -1 + 5x + 2x^2$, $q = 2 + 4x - 9x^2$
   (b) $p = x - x^2$, $q = 7 + 3x + 3x^2$
6. Let $M_{2\times 2}(\mathbb{R})$ have an inner product as : If
   $A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}$ and $B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$ then $< A, B > = a_1 b_1 + a_2 b_2 + a_3 b_3 + a_4 b_4$.
   Find the cosine of the angle between A and B.
   (a) $A = \begin{bmatrix} 2 & 6 \\ 1 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$, (b) $A = \begin{bmatrix} 2 & 4 \\ -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} -3 & 1 \\ 4 & 2 \end{bmatrix}$
7. Let $\mathbb{R}^3$ have Euclidean inner product. For which values of k are $\vec{u}$ and $\vec{v}$ orthogonal ?
   (a) $\vec{u}$ = (2, 1, 3), $\vec{v}$ = (1, 7, k), (b) $\vec{u}$ = (k, k, 1), $\vec{v}$ = (k, 5, 6).
8. In each part verify that the Cauchy-Schwarz inequality holds for the given vectors using the Euclidean inner product :
   (a) $\vec{u}$ = (1, –1), $\vec{v}$ = (–1, 4)
   (b) $\vec{u}$ = (0, –3, 1), $\vec{v}$ = (2, –1, 3)

(c) $\vec{u} = (-4, 2, 1), \vec{v} = (8, -4, -2)$

(d) $\vec{u} = (1, 1, -1, -1), \vec{v} = (1, 2, -2, 0)$

9. In each part verify that the Cauchy-Schwartz inequality holds for the given vectors :

(a) $A = \begin{bmatrix} -1 & 2 \\ 6 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 3 & 3 \end{bmatrix}$ with inner product as given in exercise 6.

(b) $p = -1 + 2x + x^2$ and $q = 2 - 4x^2$ using the inner product as given in exercise 1.

10. If V is an inner product space and if $\vec{u}$ and $\vec{v}$ are orthogonal vectors in V such that $\| \vec{u} \| = 1, \| \vec{v} \| = 2$ then show that $\| \vec{u} + \vec{v} \| = \| \vec{u} - \vec{v} \| = \sqrt{5}$.

11. Let V be an inner product space. If $\vec{w}$ is orthogonal to both $\vec{u_1}$ and $\vec{u_2}$, then show that $\vec{w}$ is orthogonal to $\alpha\vec{u_1} + \beta\vec{u_2}$, for all scalars α and β.

12. Let V be an inner product space. Then prove that the following statements are valid for all elements $\vec{u}$ and $\vec{v}$ in V :

(a) $< \vec{u}, \vec{v} > = 0$ if and only if $\| \vec{u} + \vec{v} \|^2 = \| \vec{u} \|^2 + \| \vec{v} \|^2$.

(b) $< \vec{u}, \vec{v} > = 0$ if and only if $\| \vec{u} + c\vec{v} \| \geq \| \vec{u} \|$, for all real c.

13. If $w_1, w_2, \ldots, w_n$ are any positive real numbers and $\vec{u} = (u_1, u_2, \ldots, u_n)$ and $\vec{v} = (v_1, v_2, \ldots, v_n)$ are any two vectors in $\mathbb{R}^n$ then show that :

$$| w_w u_1 v_1 + w_2 u_2 v_2 + \ldots + w_n u_n v_n | \leq (w_1 u_1^2 + \ldots + w_n u_n^2)^{1/2} (w_1 v_1^2 + \ldots + w_n v_n^2)^{1/2}.$$

14. Which of the following sets of polynomials are orthonormal with respect to the inner product as defined in exercise 1 on P_2.

(a) $\frac{2}{3} - \frac{2}{3}x + \frac{1}{3}x^2, \frac{2}{3} + \frac{1}{3}x - \frac{2}{3}x^2, \frac{1}{3} + \frac{2}{3}x + \frac{2}{3}x^2$. (b) $1, \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}x^2, x^2$.

15. Verify that the given set of vectors is orthogonal with respect to Euclidean inner product; then convert it to an orthonormal set by normalising the vectors

(a) (2, –1), (3, 6) (b) (–1, 0, 1), (2, 0, 2), (0, 5, 0)

(b) $\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right), \left(-\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{3}, \frac{1}{3}, -\frac{2}{3}\right)$.

16. If $\vec{u} = \left(\frac{1}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$ and $\vec{v} = \left(\frac{2}{\sqrt{30}}, \frac{3}{\sqrt{30}}\right)$, show that $\{\vec{u}, \vec{v}\}$ is orthonormal if $\mathbb{R}^2$ has the inner product $\langle \vec{u}, \vec{v} \rangle = 3\, u_1v_1 + 2\, u_2v_2$, but is not orthonormal if $\mathbb{R}^2$ has the Euclidean inner product.

17. Show that the vectors $\vec{v_1} = \left(-\frac{3}{5}, \frac{4}{5}, 0\right)$, $\vec{v_2} = \left(\frac{4}{5}, \frac{3}{5}, 0\right)$, $\vec{v_3} = (0, 0, 1)$ form an orthonormal basis for $\mathbb{R}^3$, with Euclidean inner product. Further, use theorem 7 to express each of the following as linear combination of $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$.

(a) $(-7, 1, 2)$, (b) $(2, -1, -2)$, (c) $\left(0, \frac{25}{12}, \frac{1}{7}\right)$.

18. Let $\vec{u_1} = \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)$, $\vec{u_2} = \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right)$, $\vec{u_3} = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$ be orthonormal basis of $\mathbb{R}^3$ with respect to Euclidean inner product. Find the co-ordinate vector of $\vec{w} = (-1, 0, 2)$.

19. Let $S = \{\vec{w_1}, \vec{w_2}, \vec{w_3}\}$, with $\vec{w_1} = \left(0, -\frac{3}{5}, \frac{4}{5}\right)$, $\vec{w_2} = (1, 0, 0)$ and $\vec{w_3} = \left(0, \frac{4}{5}, \frac{3}{5}\right)$, be given orthonormal basis for $\mathbb{R}^3$ with respect to Euclidean inner product.

(a) Find vectors u, v and w that have co-ordinate vectors

$(\vec{u})_s = (-2, 1, 2)$, $(\vec{v})_s = (3, 0, -2)$ and $(\vec{w})_s = (5, -4, 1)$.

(b) Compute $\| \vec{v} \|$, $d(\vec{u}, \vec{w})$ and $\langle \vec{w}, \vec{v} \rangle$ by using theorem 8 to co-ordinate vectors $(\vec{u})_s$, $(\vec{v})_s$ and $(\vec{w})_s$; then check the results by performing the computations directly on $\vec{u}$, $\vec{v}$ and $\vec{w}$ obtained in (a).

20. Let $\mathbb{R}^2$ have the Euclidean inner product. Use Gram-Schmidt process to transform the basis $\{\vec{u_1}, \vec{u_2}\}$ into an orthonormal basis.

(a) $\vec{u_1} = (1, -3)$, $\vec{u_2} = (2, 2)$ (b) $\vec{u_1} = (1, 0)$, $\vec{u_2} = (3, -5)$.

21. Let $\mathbb{R}^2$ have the Euclidean inner product. Use Gram-Schmidt process to transform the basis $\{u_1, u_2, u_3\}$ into an orthonormal basis :

(a) $\vec{u_1} = (1, 1, 1)$, $\vec{u_2} = (0, 1, 1)$, $\vec{u_3} = (0, 0, 1)$

(b) $\vec{u_1} = (1, 1, 1)$, $\vec{u_2} = (-1, 1, 0)$, $\vec{u_3} = (1, 2, 1)$

(c) $\vec{u_1} = (1, 0, 0), \vec{u_2} = (3, 7, -2), \vec{u_3} = (0, 4, 1)$

22. Let $\mathbb{R}^4$ have the Euclidean inner product. Use the Gram-Schmidt process to transform the basis $\{\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4}\}$ into an orthonormal basis.

$\vec{u_1} = (0, 2, 1, 0), \vec{u_2} = (1, -1, 0, 0), \vec{u_3} = (1, 2, 0, -1), \vec{u_4} = (1, 0, 0, 1)$

23. Let $\mathbb{R}^3$ have the Euclidean inner product. Find an orthonormal basis for the subspace spanned by (0, 1, 2) and (–1, 0, 1).

24. The subspace of $\mathbb{R}^3$ spanned by the vectors $\vec{u_1} = \left(\frac{4}{5}, 0, -\frac{3}{5}\right)$ and $\vec{u_2} = (0, 1, 0)$ is say W. Express $\vec{w} = (1, 2, 3)$ in the form $\vec{w} = \vec{w_1} + \vec{w_2}$, where $\vec{w_1}$ lies in W and $\vec{w_2}$ is orthonormal to W.

25. Let $\mathbb{R}^4$ have the Euclidean inner product. Express $\vec{w} = (-1, 2, 6, 0)$ in the form $\vec{w} = \vec{w_1} + \vec{w_2}$, where $\vec{w_1}$ is in space W spanned by $\vec{u_1} = (-1, 0, 1, 2)$ and $\vec{u_2} = (0, 1, 0, 1)$ and $\vec{w_2}$ is orthogonal to W.

26. Find $< \vec{u}, \vec{v} >, < \vec{v}, \vec{u} >, \| \vec{u} \|, \| \vec{v} \|$

and normalize $\vec{u}$ and $\vec{v}$, where $\vec{u} = (2, i, 3)$ and $\vec{v} = (1 + 2i, 2i, 1)$.

27. Let cos (nx) and sin (mx), where n and m are integers are functions in the vector space C ([0, 2π]). Define an inner product as

$$< f, g > = \int_0^{2\pi} f(x)\, g(x)\, dx.$$

Then show that :

(i) $\|\cos (nx)\| = \sqrt{\pi}$, if $n \neq 0$

$= \sqrt{2\pi}$, if $n = 0$,

(ii) $\|\sin (nx)\| = \sqrt{\pi}$, if $n \neq 0$

(iii) $< \cos (nx), \sin (mx) > = 0$

(iv) $< \cos (nx), \cos (mx) > = < \sin (nx), \sin (mx) > = 0$, if $m \neq n$.

Using these results, hence or otherwise find an infinite orthogonal set of vectors in C([0, 2π]).

28. Use Gram-Schmidt orthonormalization method to determine an orthonormal basis of $\mathbb{R}^3$ for given set of independent vectors :

$$\vec{u_1} = (1, 0, 1), \ \vec{u_2} = (-1, 1, 0), \ \vec{u_3} = (-3, 2, 0).$$

29. Let $f(x) = 1$, $g(x) = 1 + x$, $h(x) = x + x^2$, then f, g, h form a basis of $P_2(x)$ over $\mathbb{R}$. Define an inner product on $P_2(x)$ as :

$$< p, q > = \int_{-1}^{1} p(x)\, q(x)\, dx$$

with p(x), q(x) belong to $P_2(x)$. Construct an orthonormal basis from given set.

30. Transform the basis {(1, 0, 1), (1, 1, 0), (0, 1, 1)] to an orthonormal basis of $\mathbb{R}^3$.

Answers (6.2)

3. No. $< \vec{u}, \vec{w_2} > = -2 \neq 0$.

4. (a) $-\frac{1}{\sqrt{2}}$ (b) $-\frac{3}{\sqrt{73}}$ (c) 0 (d) $\frac{20}{9\sqrt{10}}$ (e) $-\frac{1}{\sqrt{2}}$ (f) $\frac{2}{\sqrt{55}}$

5. (a) 0 (b) 0 6. (a) $\frac{19}{10\sqrt{7}}$ (b) 0

7. (a) $k = -3$ (b) $k = -2, k = -3$

14. (a) Yes (b) No

15. (a) $\left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right), \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$, (b) $\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), (0, 1, 0)$

(c) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$

17. (a) $(-7, 1, 2) = (5)\, \vec{v_1} + (-5)\, \vec{v_2} + (2)\, \vec{v_3}$ (b) $(2, -1, -2) = (-2)\, \vec{v_1} + (1)\, \vec{v_2} + (-2)\, \vec{v_3}$

(c) $\left(0, \frac{25}{12}, \frac{1}{7}\right) = \frac{5}{3}\, \vec{v_1} + \frac{4}{5}\, \vec{v_2} + \frac{1}{7}\, \vec{v_3}$

18. $(\vec{w})_S = (0, -2, 1)$

19. (a) $\vec{u} = \left(1, \frac{14}{5}, -\frac{2}{5}\right)$, $\vec{v} = \left(0, -\frac{17}{5}, \frac{6}{5}\right)$, $\vec{w} = \left(-4, -\frac{11}{5}, \frac{23}{5}\right)$

(b) $\| \vec{v} \| = \sqrt{13}$, $d(\vec{u}, \vec{v}) = 5\sqrt{3}$, $< \vec{w}, \vec{v} > = 13$.

20. (a) $\left(\frac{1}{\sqrt{10}}, -\frac{3}{\sqrt{10}}\right), \left(\frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\right)$ (b) (1, 0) , (0, –1)

21. (a) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$

(b) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$

(c) $(1, 0, 0), \left(0, \frac{7}{\sqrt{53}}, -\frac{2}{\sqrt{53}}\right), \left(0, \frac{2}{\sqrt{53}}, \frac{7}{\sqrt{53}}\right)$

22. $\left(0, \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0\right), \left(\frac{5}{\sqrt{30}}, -\frac{1}{\sqrt{30}}, \frac{2}{\sqrt{30}}, 0\right)$

$\left(\frac{1}{\sqrt{10}}, \frac{1}{\sqrt{10}}, -\frac{2}{\sqrt{5}}, -\frac{2}{\sqrt{10}}\right), \left(\frac{1}{\sqrt{15}}, \frac{1}{\sqrt{15}}, -\frac{2}{\sqrt{15}}, \frac{3}{\sqrt{15}}\right)$

23. $\left(0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right), \left(-\frac{\sqrt{5}}{\sqrt{6}}, -\frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}}\right)$.

24. $\vec{w_1} = \left(-\frac{4}{5}, 2, \frac{3}{5}\right), \vec{w_2} = \left(\frac{9}{5}, 0, \frac{12}{5}\right)$

25. $\vec{w_1} = \left(-\frac{5}{4}, -\frac{1}{4}, \frac{5}{4}, \frac{9}{4}\right), \vec{w_2} = \left(\frac{1}{4}, \frac{9}{4}, \frac{19}{4}, -\frac{9}{4}\right)$.

26. $< \vec{u}, \vec{v} > = 7 - 4i, \quad < \vec{v}, \vec{u} > = 7 + 4i, \quad \| \vec{u} \| = \sqrt{14}, \quad \| \vec{v} \| = \sqrt{10}.$

$\frac{\vec{u}}{\| \vec{u} \|} = \frac{1}{\sqrt{14}} (2, i, 3), \quad \frac{\vec{v}}{\| \vec{v} \|} = \frac{1}{\sqrt{10}} (1 + 2i, 2i, 1).$

27. $\left\{\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \ldots, \frac{\cos (nx)}{\sqrt{\pi}}, \ldots, \frac{\sin x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}, \ldots\right\}$

is an infinite orthonormal set of vectors in C ([0, 2π]).

28. $\vec{e_1} = \frac{1}{\sqrt{2}} (1, 0, 1), \quad \vec{e_2} = \frac{1}{\sqrt{6}} = (-1, 2, 1), \quad \vec{e_3} = \frac{1}{\sqrt{3}} (-1, -1, 1).$

29. $\left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{6}}{2} x, \frac{\sqrt{10}}{4} (3x^2 - 1)\right\}$ is orthonormal basis for $P_2(x)$.

30. $\left\{\frac{1}{\sqrt{2}} (1, 0, 1), \frac{1}{\sqrt{6}} (1, 2, -1), \frac{1}{\sqrt{3}} (-1, 1, 1)\right\}$ orthonormal basis for $\mathbb{R}^3$.

Important Points

- Definition of inner product on a vector space and its properties, examples.
- Definition of inner product space and examples.
- Schwartz inequality in $\mathbb{C}^n$ and $\mathbb{R}^n$.
- Norm of a vector, distance between two vectors and their properties.
- Orthogonal and orthonormal set of vectors with examples.
- Gram-Schmidt orthonormalization process and examples.
- Orthogonal projections and examples.
- Rank and nullity of a matrix.
- Sylvester inequality and its significance.

Miscellaneous Exericse

(A) Theory Questions :

1. Define :
   (a) Inner product on a vector space.
   (b) Inner product space.
   Illustrate each by an example.
2. State and prove the consequences of an inner product.
3. State and prove Schwartz inequality.
4. Define : (a) norm of a vector, (b) distance between two vectors, (c) orthogonal set and (d) orthonormal set. Illustrate each by an example.
5. Prove that the set of non-zero vectors which is orthogonal is linearly independent.
6. Define orthogonal projection and illustrate it.
7. Explain Gram-Schmidt process for orthonormalization.
8. State and prove Sylvester inequality.

(B) Numerical Problem :

1. Show that the set of vectors $\vec{u_1} = (1, i, 0)$, $\vec{u_2} = (i, 1, 0)$ and $\vec{u_3} = (0, 0, 2 + 3i)$ is an orthogonal set. Further transform this set to orthonormal set.
2. Find $< \vec{u}, \vec{v} >, < \vec{v}, \vec{u} >$, $\| \vec{u} \|, \| \vec{v} \|$ and the normalised vectors of $\vec{u}$ and $\vec{v}$, where $\vec{u} = (2, i, 3)$, $\vec{v} = (1 + 2i, 2i, 1)$.
3. Check whether the vectors $\vec{u} = (i - 1, i, 1)$ and $\vec{v} = (1, i + 1, 0)$ are linearly independent. If linearly independent, construct on orthonormal set of vectors with the help of them.
4. Do the same exercise as in (3),
   where $\vec{u} = (1, 2 + i, i, 1 - i)$, $\vec{v} = (1 + i, 2 + i, - 5, - 6)$.
5. Resolve the vector $\vec{v}$ into two perpendicular components, one alone $\vec{u}$, where $\vec{u} = (1, 2 + i, i, 1 - i)$, $\vec{v} = (1 + i, 2 + i, - 5, - 6)$.
6. Find an orthonormal basis in $\mathbb{R}^3$ using the Gram-Schemidt process :
   $$(i, 0, 1), (i - 1, i, 1), (1, i + 1, 0).$$

Answers

2. $< \vec{u}, \vec{v} > = 7 - 4i$, $< \vec{v}, \vec{u} > = 7 + 4i$, $\| \vec{u} \| = \sqrt{14}$, $\| \vec{v} \| = \sqrt{10}$. Normalised vector are :
   $\left(\frac{2}{\sqrt{14}}, \frac{i}{\sqrt{14}}, \frac{3}{\sqrt{14}} \right)$, $\left(\frac{1+2i}{\sqrt{10}}, \frac{2i}{\sqrt{10}}, \frac{1}{\sqrt{10}} \right)$.
3. $\frac{1}{2}(i - 1, i, 1)$, $\frac{1}{2\sqrt{2}}(1 - i, 1 + 2i, i)$.
4. $\frac{1}{3}(1, 2 + i, i, 1 - i)$, $\frac{1}{\sqrt{68}}(1 + i, 2 + i, -5, -6)$.
5. $(0, 0, 0, 0)$, $(1 + i, 2 + i, -5, -6)$.
6. $\frac{1}{\sqrt{2}}(i, 0, 1)$, $\frac{1}{\sqrt{6}}(1, -2i, i)$, $\frac{1}{\sqrt{15}}(2 - i, 1 + 2i, 1 + 2i)$.

❑❑❑

Chapter 7...

Quadratic Forms

Synopsis

7.1 Introduction of Quadratic Form
7.2 Singular and Non-singular Linear Transformations
7.3 Linear Transformations of Quadratic Form on Field F
7.4 Congruence of Quadratic Forms and Matrices
7.5 Elementary Congruent Transformations of Non-singular Matrix
7.6 Elementary Congruent Transformations
7.7 Congruent Reduction of a Symmetric Matrices
7.8 Congruence Reduction of Skew-symmetric Matrices
7.9 Properties of Definite, Semi-definite and Indefinite Forms

Objectives

At the end of this chapter you will be able to understand :

(i) What is quadratic form and the nature of a quadratic form ?

(ii) How to reduce a symmetric matrix into a quadratic form ?

(iii) The property of definite matrices.

Vasant Shankar Huzurbazar

V.S.Huzurbazar was born on 15th September,1919 at Kolhapur in Maharashtra State, INDIA. As a student he had a brilliant career in University of Bombay, Banaras Hindu University and University of Cambridge, England where he obtained his Ph.D. degree for his thesis entitled "Properties of Sufficient Statistics" written under the supervision of Sir Harold Jeffreys.

Professor Huzurbazar is internationally well-known for his work on maximum likelihood estimation, invariants for probability distributions and sufficient statistics.

During his long career, Professor Huzurbazar has received many recognitions and Honours. He played an active role in the development of Statistics and Mathematics in India. The Government of India recognized his contributions to education by awarding him the "Padma Bhushan" in 1974.

After his return from Cambridge in 1949, Dr. Huzurbazar worked in the University of Gauhati, Lucknow and also in the Bureau of Economics and Statistics Of Government of Bombay. He Joined the University of Pune as Professor and Head of the Department of Mathematics and Statistics in 1953.

7.1 Introduction of Quadratic Form

In linear algebra the linear equations of the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

The expression of the form $a_1x_1 + a_2x_2 + \ldots + a_nx_n$ is a function of n variables, called a **linear form**. In linear form all variables occur to the first power, and there is product of variables in the expression.

In this section we will be concerned with **quadratic forms**, which are the functions of the form.

$$a_1x_1^2 + a_2x_2^2 + \ldots a_nx_n^2 + (\text{all possible terms of the form } a_k\, x_i\, x_j \text{ for } i < j). \quad \ldots (7.1)$$

For example, the most general quadratic form in two variables x_1 and x_2 is

$$a_1x_1^2 + a_2x_2^2 + a_3x_1x_2 \quad \ldots (7.2)$$

The most general quadratic form in three variables x_1, x_2 and x_3 is

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + a_4x_1x_2 + a_5x_1x_3 + a_6x_2x_3 \quad \ldots (7.3)$$

The terms in the quadratic form that involve products of distinct variables are called the **cross-product terms**. Thus in (7.2) the last term is cross-product term and in (7.3) the last three terms are cross-product terms.

We can write equation (7.2) in matrix form as

$$[x_1 \;\; x_2] \begin{bmatrix} a_1 & \frac{a_3}{2} \\ \frac{a_3}{2} & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \ldots (7.4)$$

and equation (7.3) can be written as;

$$[x_1 \;\; x_2 \;\; x_3] \begin{bmatrix} a_1 & \frac{a_4}{2} & \frac{a_5}{2} \\ \frac{a_4}{2} & a_2 & \frac{a_6}{2} \\ \frac{a_5}{2} & \frac{a_6}{2} & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \ldots (7.5)$$

The products in equation (7.4) and (7.5) are both the form $\mathbf{X^T A X}$**, where, X is the column vector of variables and A is a symmetric matrix whose diagonal entries are the coefficients of the squared terms and whose entries off the main diagonal are half the coefficients of the cross-product terms.** More precisely, the diagonal entry in row i is the coefficient of x_i^2 and the off diagonal entry in row i and column j is half the coefficient of the product x_ix_j.

SOLVED EXAMPLE

Example 7.1 : (Matrix Representation of Quadratic Form)

Represent the following quadratic forms in matrix form.

(a) $4x^2 - 6xy + 3y^2$

(b) $8x^2 - 5y^2$

(c) $3xy$

(d) $2x_1^2 - 5x_2^2 + 3x_3^2 + 6x_1x_2 - 3x_1x_3 + 8x_2x_3$.

Solution : Using the above discussion and (4) and (5) we have,

(a) $$4x^2 - 6xy + 3y^2 = [x \quad y] \begin{bmatrix} 4 & -3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(b) $$8x^2 - 5y^2 = [x \quad y] \begin{bmatrix} 8 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(c) $$3xy = [x \quad y] \begin{bmatrix} 0 & \frac{3}{2} \\ \frac{3}{2} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(d) $2x_1^2 - 5x_2^2 + 3x_3^2 + 6x_1x_2 - 3x_1x_3 + 8x_2x_3$.

$$= [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 2 & 3 & -\frac{3}{2} \\ 3 & -5 & 4 \\ -\frac{3}{2} & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Note :

(1) Symmetric matrices are useful, but not essential, for representing quadratic forms. For example, the quadratic form $x^2 + 3y^2 - 4xy$ has the symmetric matrix representation as

$$x^2 + 3y^2 - 4xy = [x \quad y] \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

However it can also be written as

$$x^2 + 3y^2 - 4xy = [x \quad y] \begin{bmatrix} 1 & -1 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

(2) However symmetric matrices are usually more convenient to work with, so when we write the quadratic form as $X^T A X$, it will be always understood even if it is not stated explicitly, that **A is symmetric matrix**.

In this section we shall study **only** the following two problems, which involve quadratic forms.

(i) To find the maximum and minimum values of the quadratic form $X^T A X$ if X is constrained so that

$$\|X\| = (x_1^2 + x_2^2 + \ldots x_n^2)^{1/2} = 1.$$

(ii) What conditions must A satisfy in order for a quadratic form to satisfy the inequality $X^T A X > 0$ for all $X \neq 0$?

We state only the following theorem where proof is not expected. This theorem gives answer to the two problems (i) and (ii).

Theorem 1 : Let A be a symmetric $n \times n$ matrix with eigenvalues $\lambda_n \leq \lambda_{n-1} \leq \ldots \leq \lambda_1$ and $\|X\| = 1$, then

(a) $\lambda_1 \geq X^T A X \geq \lambda_n$

(b) $X^T A X = \lambda_n$, if X is an eigenvector of A corresponding to λ_n and $X^T A X = \lambda_1$ if X is an eigenvector of A corresponding to λ_1.

Remark : From above theorem it follows that subject to the constraint

$$\|x\| = (x_1^2 + x_2^2 + \ldots x_n^2)^{1/2}) = 1.$$

(i) The part (a) of the theorem asserts that the quadratic form $X^T A X$ has **maximum value λ_1** (the largest eigenvalue) and **minimum value λ_n** (the smallest eigenvalue).

(ii) The part (b) of the theorem states that the quadratic form $X^T A X$ **attains its maximum** at the eigenvector X corresponding to the eigenvalue λ_1 and it **attains its minimum** at the eigenvector X corresponding to the eigenvalue λ_n.

SOLVED EXAMPLES

Example 7.2 : Find the maximum and the minimum values of each of the quadratic form subject to the constraint $x^2 + y^2 = 1$ and determine the values of x and y at which the maximum and minimum occur.

(a) $5x^2 - y^2$ (b) $2x^2 + y^2 + 3xy$.

Solution : (a) The matrix form of the quadratic form $5x^2 - y^2$ is

$$[x \quad y] \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

So hence $A = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$, we require to find eigenvalues of A for first part and for second part eigenvectors corresponding to the eigenvalues of A.

Clearly the eigenvalues of A are 5 and – 1. Therefore, 5 is the maximum value and – 1 is the minimum value of the given quadratic form.

Now to find eigenvector X corresponding to $\lambda = 5$, we have to solve $(\lambda I - A) X = 0$, with $\lambda = 5$, which reduces to

$$\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This implies that $x = t$, $t \in R$ and $y = 0$. So that $X = (1, 0)$ is the eigenvector of A corresponding to $\lambda = 5$. Thus, the maximum value of the quadratic form is 5 at $x = 1$ and $y = 0$.

To find the eigenvector X corresponding to $\lambda = -1$, we solve $(\lambda I - A)\, X = 0$ with $\lambda = -1$.

This reduces solving

$$\begin{bmatrix} -6 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives $x = 0$ and $y = 1$, hence minimum value -1 occurs at $x = 0$ and $y = 1$.

(b) The matrix form of the quadratic form is

$$[x \;\; y] \begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Thus, $$A = \begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}$$

The eigenvalues of A are obtained by solving $\det(\lambda I - A) = 0$. This gives $\lambda_1 = \frac{3 + \sqrt{10}}{2}$ and $\lambda_2 = \frac{3 - \sqrt{10}}{2}$. The maximum value of quadratic form is $\frac{3 + \sqrt{10}}{2}$ and minimum is $\frac{3 - \sqrt{10}}{2}$. Solving the system $(\lambda I - A)\, X = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for λ_1 and λ_2, we obtain $x = \pm \frac{3}{\sqrt{20 - 2\sqrt{10}}}$, $y = \pm \frac{3}{\sqrt{20 + 2\sqrt{10}}}$ at which maximum value occurs. And $x = \pm \frac{3}{\sqrt{20 + 2\sqrt{10}}}$ and $y = \pm \frac{3}{\sqrt{20 - 2\sqrt{10}}}$ at which minimum value occurs.

Example 7.3 : Find the maximum and minimum value of the quadratic form a subject to the constraint $x_1^2 + y_2^2 + z_3^2 = 1$, where quadratic form

$$Q(x_1, x_2, x_3) = 3x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_3.$$

Solution : The matrix form of the quadratic form

$$[x_1 \;\; x_2 \;\; x_3] \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \text{ where, } A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}.$$

We have to find eigenvalues of A and the corresponding eigenvectors. To find eigenvalues we have to find det $(\lambda I - A) = 0$. We have

$$\det(\lambda I - A) = \det\begin{bmatrix} \lambda - 3 & 0 & -1 \\ 0 & \lambda - 2 & 0 \\ -1 & 0 & \lambda - 3 \end{bmatrix}$$

$$= (\lambda - 3)(\lambda - 2)(\lambda - 3) - 1(\lambda - 2) = (\lambda - 2)^3 (\lambda - 4)$$

$\therefore$ $\lambda_1 = 4$ and $\lambda_2 = 2$ are the eigenvalues of A. Thus the quadratic form has maximum value 4 and minimum 2. Now to find the eigenvectors of A corresponding to $\lambda_1 = 4$ and $\lambda_2 = 2$, we have to solve the system of linear equations $(\lambda I - A) X = 0$, where,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

For $\lambda_1 = 4$, we get $X = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

We have to normalise X, so that $X_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ is the eigenvector of A corresponding to $\lambda_1 = 4$, at which the quadratic form has maximum value 4. Similarly, for $\lambda_2 = 2$, we get $X = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $X = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Normalising these vectors. We see that the quadratic form has minimum value 2 at each of the eigenvectors $X = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$ and $X = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Definition : A quadratic form X^TAX is called **positive definite** if $X^TAX > 0$ for all $X \neq \vec{0}$, and a symmetric matrix A is called a **positive definite matrix** if X^TAX is a positive definite quadratic form.

The following theorem states when a symmetric matrix is positive definite matrix.

Theorem 2 : A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive.

Example 7.4 : Check whether the symmetric matrices

(a) $\begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$ are positive definite.

Solution : (a) The eigen values of the matrix $\begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$ are [5 and – 1]. Since they are not all positive therefore the matrix is **not** positive definite.

(b) The matrix $\begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}$ has eigenvalues $\frac{3+\sqrt{10}}{2}$ and $\frac{3-\sqrt{10}}{2}$, which are not all positive, the matrix is **not** positive definite in this case also.

(c) Here 2, 2 and 4 are the eigenvalues of the matrix $\begin{bmatrix} 3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$ hence this matrix is positive definite.

Another criterion for showing a symmetric matrix is positive definite.

Definition : If

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ : & : & & : \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

is a square matrix, then the **principal submatrices** of A are the submatrices formed from the first k rows and k columns of A for k = 1, 2, … n. These submatrices are

$$A_1 = [a_{11}],\ A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\ A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},\ \dots\ A_n = A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ : & : & & : \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Theorem 3 : A symmetric matrix A is positive definite if and only if the determinant of every principal submatrix is positive.

Example 7.5 : Determine whether the matrix A is positive definite, where, $A = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$.

Solution : The principal submatrices of A are $A_1 = [3]$, $A_2 = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$ and $A_3 = A$.

We see that : $\det A_1 = 3$, $\det A_2 = 5$ and $\det A_3 = \det A = 12$, so the determinants of all principal submatrices of A are positive, hence A is positive definite and the quadratic form $X^TAX > 0$ for all $X \neq 0$, that is, positive definite.

Remark : A symmetric matrix A and the quadratic form $X^T A X$ are called.

positive semidefinite	if	$X^T A X \geq 0$ for all X
negative definite	if	$X^T A X < 0$ for all $X \neq 0$
negative semidefinite	if	$X^T A X \leq 0$ for all X
indefinite	if	$X^T A X$ has both positive and negative values.

Note : A symmetric matrix A is **positive semidefinite** if and only if all of its eigenvalues are nonnegative. Also, A is positive semidefinite if and only if all its principal submatrices have non-negative determinants.

Exercise 7.1

1. Which of the following are quadratic forms ?
   (a) $x^2 + \sqrt{3}\, y^2$ (b) $x^2 - 2y^3 + 4xz$
   (c) $4x^2 - 2y^2 + z^2 - 8x_1x_2$ (d) $x^2 + y^2 + z^2 + xyz$
   (e) $xy + 5yz + 6xz$ (f) $x^2 + y^2 - 6z^2 + x + y - z$
   (g) $(x + 3z)^2$ (h) $(x + y)^2 - 2(x - z)^2$
2. Express the following forms in matrix notation $X^T A X$, where A is a symmetric matrix.
   (a) $3x^2 - 9y^2$ (b) $3x^2 - 5y^2 + 6xy$
   (c) $3x^2 + 4x_1x_2$ (d) $7xy$
3. Express the following quadratic forms in the matrix notation X^TAX, where A is a symmetric matrix
   (a) $6x^2 - y^2 + 5z^2 - xy + 4yz + 3xz$ (b) $x^2 - y^2 - z^2 + 6xy - 4xz$
   (c) $xy + yz$ (d) $\sqrt{2}\, x^2 + \sqrt{5}\, y^2 - \sqrt{6}\, z^2 + 8\sqrt{5}\, xz$
   (e) $x_1^2 + x_2^2 - x_4^2 + x_1x_2 + x_2x_3 + 3x_1x_4$

4. In each of the following, find a formula for the quadratic form that does not use matrices.

   (a) $[x \;\; y]\begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$ (b) $[x_1 \;\; x_2]\begin{bmatrix} 5 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

   (c) $[x \;\; y \;\; z]\begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & -2 \\ 0 & -2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

5. In each of the following, find the maximum and minimum values of the quadratic form subject to the constraint $x^2 + y^2 = 1$ and determine the values x and y at which maximum and minimum occur.

   (a) $x^2 + y^2 + 4x_1x_2$ (b) $5x^2 + 2y^2 - xy$

6. Which of the following matrices are positive definite.

   (a) $\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -2 \\ -2 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 5 & 0 \\ 0 & -3 \end{bmatrix}$

7. Determine which of the matrices are in exercise (6) are positive semi definite (or non-negative definite).

8. Determine which of the following matrices are positive definite ?

   (a) $\begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ (d) $\begin{bmatrix} -5 & 1 & 2 \\ 1 & 3 & 0 \\ 2 & 0 & 4 \end{bmatrix}$

9. Determine which of the matrices in exercise (8) are positive definite.

Answers (7.1)

1. (a), (c), (e), (g) and (h) are quadratic forms.

2. (a) $\begin{bmatrix} 3 & 0 \\ 0 & -9 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 3 \\ 3 & -5 \end{bmatrix}$

   (c) $\begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & \frac{7}{2} \\ \frac{7}{2} & 0 \end{bmatrix}$

3. (a) $\begin{bmatrix} 6 & -\frac{1}{2} & \frac{3}{2} \\ -\frac{1}{2} & -1 & 2 \\ \frac{3}{2} & 2 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 3 & -2 \\ 3 & -1 & 0 \\ -2 & 0 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$ (d) $\begin{bmatrix} \sqrt{2} & 0 & 4\sqrt{5} \\ 0 & \sqrt{5} & 0 \\ 4\sqrt{5} & 0 & -\sqrt{6} \end{bmatrix}$ (e) $\begin{bmatrix} 1 & \frac{1}{2} & 0 & \frac{3}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ \frac{3}{2} & 0 & 0 & -1 \end{bmatrix}$

4. (a) $x^2 + 4y^2 - 4xy$ (b $5x_1^2 + x_2^2 + 3x_1x_2$ (c) $x^2 + 5y^2 - z^2 + 4xy - 4yz$

5. (a) maximum 3 at $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ minimum -1 at $x = \frac{1}{\sqrt{2}}$, $y = \frac{1}{\sqrt{2}}$.

(b) maximum value $\frac{7+\sqrt{10}}{2}$ at $\pm\left(\frac{1}{\sqrt{20-6\sqrt{10}}}, \frac{-1}{\sqrt{20+6\sqrt{10}}}\right)$

minimum value $\frac{7-10}{2}$ at $\pm\left(\frac{1}{\sqrt{20+6\sqrt{10}}}, \frac{1}{\sqrt{20-6\sqrt{10}}}\right)$

6. (b) 8. (a) 9. (a)

7.2 Singular and Non-Singular Linear Transformations

The linear transformation to be now considered will be in the context of the quadratic forms. Let

$$x_1, \ldots, x_n \text{ and } y_1, \ldots, y_n$$

be two sets of variables. Consider, now, a system of linear equations

$$\left.\begin{aligned} y_1 &= p_{11}x_1 + p_{12}x_2 + \ldots + p_{1n}x_n \\ y_2 &= p_{21}x_1 + p_{22}x_2 + \ldots + p_{2n}x_n \\ \ldots \quad &\ldots \quad \ldots \quad \ldots \quad \ldots \\ \ldots \quad &\ldots \quad \ldots \quad \ldots \quad \ldots \\ y_n &= p_{n1}x_1 + p_{n2}x_2 + \ldots + p_{nn}x_n \end{aligned}\right\} \quad \ldots (1)$$

expressing each of $y_1, \ldots, y_n$ in terms of $x_1, \ldots, x_n$.

This system of liner equations is equivalent to the single matrix equation

$$Y = PX \quad \ldots (2)$$

where, $P = \begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ p_{1n} & p_{n2} & \ldots & p_{nn} \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ \ldots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ \ldots \\ \ldots \\ y_n \end{bmatrix}$

The system (1) of linear equations of the equivalent matrix equation (2) assign to *each n-vector X, a n-vector Y.* Suppose, now, that the matrix P is *non-singular*, so that we may rewrite (2) as

$$X = P^{-1}Y \qquad \ldots (3)$$

Thus equation shows that the relation between the pairs of n-vectors X, Y is one-to-one, for each Y is the correspondent of *one and only one X.*

We say that the system (1) or the equivalent matrix equation (2) determines a *linear transformation* and that the linear transformation is non-singular or singular according as the matrix P of the linear transformation is non-singular or singular.

Hence, we notice that each linear transformation is associated with a matrix, so that a matrix, now, appears as representative of linear transformation.

Exercises 7.2

1. Give the matrices corresponding to the following linear transformations :

(i) $y_1 = \sqrt{2}x_1 + x_2$
$y_2 = 3x_1 - 3\sqrt{3}x_2$

(ii) $y_1 = x_1 - 2x_2 + 3x_3$
$y_2 = 2x_1 + x_2$
$y_3 = 3x_1 + 4x_2 - 5x_3$

(iii) $y_1 = 5x_1 + x_2 - x_3$
$y_2 = x_1 - 5x_2 + x_3$
$y_3 = x_1 + x_2 - 5x_3$

(iv) $y_1 = ix_1 - \sqrt{2}x_2$
$y_2 = \sqrt{3}x_1 + (\sqrt{2} + i)\, x_2$

2. Give the linear transformation corresponding to the following matrices :

$$\begin{bmatrix} 1 & -2 & -5 \\ 0 & 7 & -1 \\ 2 & 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & -2 & 1 \\ 3 & -1 & -3 \\ -4 & 4 & 5 \end{bmatrix}; \quad \begin{bmatrix} \sqrt{2} & \sqrt{3} \\ \sqrt{5} & -\sqrt{7} \end{bmatrix} \cdot \begin{bmatrix} i & 2i \\ 0 & 2-\sqrt{2i} \end{bmatrix}.$$

7.2.1 Linear Transformations

Consider, now, the two linear transformations of V_n defined by the matrix equations

$$Y = PX, \; Z = QY$$

By virtue of these matrix equations, to each vector X corresponds a vector Y and to each vector Y corresponds a vector Z so that these two equations, taken together, associate a vector Z to each vector X.

The resultant transformation from X to Z determined by the single equation

$$Z = Q(PX) = QPX$$

is linear and corresponds to the product QP of the matrices Q and P which separately corresponds to the given linear transformations. Thus *the matrix of the resultant of two linear transformations is the product of the matrices of their transformations.*

It is also obvious that the *Resultant of two non-singular transformations is itself non-singular.*

Example 7.6 : Considering the linear transformations :

$$\begin{bmatrix} y_1 = 2x_1 - 3x_2 + 4x_3 \\ y_2 = x_1 + x_2 + x_3 \\ y_3 = 3x_1 - 2x_2 + 4x_3 \end{bmatrix} \cdot \begin{bmatrix} z_1 = y_1 + y_2 - y_3 \\ z_2 = 2y_1 - y_2 \\ z_2 = y_1 - y_2 + 2y_3 \end{bmatrix}$$

verify the result arrived at above for the matrix of the resultant of linear transformations.

7.3 Linear Transformation of Quadratic Form on Field F

In case the coefficients of a quadratic form all belong to the field F we say that the form belongs to the field F. In particular if the coefficient of a quadratic form are all real, we say that the quadratic form is real.

Similarly, we shall be considering linear transformations from the vector space $V_n(F)$ of n-tuples of members of the field F to itself.

Consider a quadratic form

$$X'AX \quad \ldots \text{(i)}$$

A being a symmetric matrix over the field F. Also consider a *non-singular* linear transformation given by;

$$X = PY \quad \ldots \text{(ii)}$$

P being an non-singular matrix over the field F.

Transposing (ii), we get

$$X' = (PY)' = Y'P'$$

Thus we have, $X'AX = Y'P'APY = Y'By$, say

Thus, we have

where, $B = P'AP$

We notice that B is also a symmetric matrix over the field F. In fact, we have

$$B = (P'AP)' = P'A'(P')' = P'AP = B$$

The quadratic form Y'BY which is also a quadratic form over the field F' is called a *linear transform* of the quadratic form X'AX. Thus if a quadratic form with matrix A over a field F is subjected to a linear transformation with matrix P over F, then the transform is a quadratic form over F with matrix P'AP.

7.4 Congruence of Quadric Forms and Matrices

Definition : An n-rowed square matrix B on field F is said to be congruent to another n-rowed square matrix A on field F, if there exists a non-singular matrix P on field F such that

$$B = P'AP$$

The "*Congruence of matrices*" has the following properties :

(i) Reflexivity : If A is conguent to B, then B is also congruent to A, for

$$A = P'BP$$

$$\Rightarrow \quad B = (P')AP^{-1} = (P^{-1})'\, AP^{-1}$$

(ii) Transitivity : If A be congruent to B and B to C, then A is congruent to C, for

$$\left.\begin{matrix} A = P'BP \\ B = Q'CQ \end{matrix}\right\} \Rightarrow A = P'Q'CQP = (QP'\, C(QP)$$

Because of the property (ii) of symmetry, we may say that two matrices are congruent instead of saying that one matrix is congruent to the other.

Thus, the relation "Congruent over F' is reflxive, symmetric and transitive'.

Also, we say that two quadratic forms over a field F are congruent on field F, i.e., if then corresponding symmetric matrices are congruent over F, i.e., if there exists a linear non-singular transformation which transforms one form to the other.

Theorem 4 : The ranges of values of two congruent quadratic forms are the same.

Let $X'AX = Y'BY$

be two quadratic congruent forms so that there exists a non-singular matrix P such that

$$B = P'AP$$

We have to show that as the vectors X and Y range over the space $V_n(F)$, the sets of values assumed by the two forms are the same. We write

$$X = PY \Leftrightarrow Y = P^{-1}X$$

Consider any value $X_1'AX_1$ of the first form for any vector X_1 of V_n.

Let $P^{-1}X_1 = Y_1$

Now, $Y_1'BY_1 = (P^{-1}X_1)'\, P'AP\, (P^{-1}X_1)$

$$= X_1'\, (P')^{-1}\, P'\, APP^{-1}\, X_1 = X_1'\, AX_1$$

so that each value of X'AX is a value of Y'BY.

7.5 Elementary Congruent Transformations of a Non-singular Matrix

Let B be congruent to A so that there exists a non-singular matrix P such that

$$B = P'AP$$

As every non-singular matrix is a product of elementary matrices. We write

$$P = P_1P_2 \ldots P_{s-1}\, P_s$$

where, $P_1, \ldots, P_s$ are all elementary matrices. We have

$$P' = P'P_{s-1}' \ldots P_2'P_1' \qquad \text{(by reversal law of transpose)}$$

where, $P_1', \ldots, P_i'$

Also elementary matrices.

We have, $P'AP = P_s' P_1' A P_1 P_2 \dots P_{s-1} P_s$

$= P_s' [P_{s-1}' \{\dots P_2' \{P_1' AP_1\} P_2 \dots\} P_{s-1}] P_s \quad \dots (1)$

Now, $P_1'AP_1$ is obtained from A by subjecting the latter to a pair of elementary transformations, one row and the other column such that of two corresponding elementary matrices P_1', P_1, either is the transpose of the other. It will thus appear from (1) that P'AP arises by subjecting A to a series of such pairs of elementary row transformations.

7.6 Elementary Congruent Transformations

Definition : A pair of elementary transformations, one row and the other column is said to constitute one elementary congruent transformation if the corresponding elementary matrices are such that each is the transpose of the other.

The following are the three types of elementary transformations :

I. Interchange of the i^{th} and j^{th} row as well as of the i^{th} and j^{th} column.

II. Multiplication of the i^{th} row as well as the j^{th} column by a non-zero number c.

III. Addition of k times the j^{th} row to the i^{th} row and of k times the j^{th} column to the i^{th} column.

Every matrix obtained from a given matrix A by subjecting it to a series of elementary congruent transformations is congruent to A.

It may be seen that the transforms of a matrix A by the three elementary congrent transformations above are

$$E_{ij}(A),\ E_{ij},\ E_i(c),\ A\,E_i(c),\ E_{ij}(k),\ A\,E_{ij}(k)$$

respectively, for

$$E_{ij} = E_{ij},\ E_i'(c) = E_i(c),\ E_{ij}'(k) = E_{ij}(k).$$

7.7 Congruent Reduction of a Symmetric Matrices

Theorem 5 : If, V, be any n-rowed non-zero symmetric matrix of rank, r over a field F, then there exists an n-rowed non-singular matrix P over F such that

$$P'AP = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where, A_1, is a non-singular r-rowed diagonal matrix over F and each 0, is a zero matrix of an appropriate type.

It has to be shown that, every symmetric matrix A of rank r over F is congruent over F to a diagonal matrix having just r non-zero diagonal elements.

Firstly, we show that, given a non-zero symmetric matrix A over F, then there exists a matrix B over F congruent to A such that $b_{11} \neq 0$.

Case I : In case $a_{11} \neq 0$, we take B the same as A.

Case II : If $a_{11} = 0$ but some diagonal element, say $a_{ij} \neq 0$, then, by interchanging the i^{th} row with first row and the i^{th} column with the first column, we obtain a matrix B congruent to A such that $b_{11} = a_{11} \neq 0$.

Case III : Suppose, now, that every diagonal element of the matrix A is 0. As A is a zero matrix, there exists some non-zero element of A. Let it be a_{ij}. Thus, $a_{ij} = a_{ji} \neq 0$.

Then by adding the j^{th} row to the i^{th} row and the j^{th} column to the i^{th} column, we obtain a matrix C congruent to A such that

$$c_{ij} = a_{ij} + a_{ji} = 2a_{ij} \neq 0$$

and then by interchanging the i^{th} row and the i^{th} column of C with its first row and first column respectively, we obtain a matrix B congruent to C and therefore, also congruent to A such that

$$b_{11} = c_{jj} \neq 0$$

Consider, now, a matrix B congruent to the symmetric matrix A such that $b_{11} \neq 0$. Being congruent to a symmetric matrix, B is itself symmetric.

Let $$k_p = b_{11}^{-1} b_{1p} = b_{11}^{-1} b_{p1}$$

so that k_p belongs to F.

Subtracting from the p^{th} row, k_p times the first row and from the p^{th} column, k_p times the first column, we obtain a matrix, D congruent to the symmetric matrix B such that

$$d_{1p} = 0, \; d_{p1} = 0$$

As in the preceding step, subtracting from each of the rows and columns, suitable multiples of the first row and the first column, we obtain a matrix E congruent to A such that each member of the first row and first column excepting the leading element is zero. Thus the matrix E is of the form

$$\begin{bmatrix} b_{11} & 0 \\ 0 & E_1 \end{bmatrix}$$

where, E_1 is an $(n - 1)$ rowed symmetric matrix.

If E_1 be not a zero matrix then we can deal with E_1 as we did with A without affecting the first row and the first column of E.

Thus, proceeding, we shall obtain a diagonal matrix congruent to A. As congruence is a special case of equivalence, the rank must remain unaltered so that the finally obtained diagonal matrix must have just, r, non-zero diagonal elements.

At each stage of the above process of reduction, we have employed an elementary congruent transformation. Since each elementary congruent transformation can be effected by pre-multiplication and post-multiplication with a pair of elementary matrices each of which is the transpose of the other, we see that there exists a non-singular matrix P, viz., the product of the post-multiplied elementary matrices such that P'AP is a diagonal matrix with just r, non-zero diagonal elements.

As also the elements of the elementary matrices employed are obtained by performing only rational operations on those of A, we see that the matrices P, P' and P'AP all belong to the field F of the matrix A.

SOLVED EXAMPLES

Example 7.6 : Consider matrix A which is symmetric and reduce in to a fundamental form and interpret the result of quadratic forms :

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

Solution : We wrtie, A = I'AI

$$\Leftrightarrow \quad \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Performing the two elementary congruent transformations,

$$R_{21}(1/3), \quad C_{21}(1/3)$$

$$R_{31}(-1/3), \; C_{31}(-1/3)$$

We obtain the equality

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 1/3 & -1/3 \\ 0 & -1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1/3 & 1 & 0 \\ -1/3 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 1/3 & -1/3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Now use the elementary congruent transformation, we get

$$R_{12}(1/7), C_{12}(1/7)$$

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 7/3 & 0 \\ 0 & 0 & 19/7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ -2/7 & 1/7 & 0 \end{bmatrix} A \begin{bmatrix} 1 & 1/3 & -2/7 \\ 0 & 1 & 1/7 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, we obtain the diagonal matrix

$$B = \text{Diag.} \left(6, \frac{7}{3}, \frac{16}{7}\right)$$

which is congruent to the given symmetric matrix. Here

$$P = \begin{bmatrix} 1 & 1/3 & -2/7 \\ 0 & 1 & 1/7 \\ 0 & 0 & 1 \end{bmatrix}$$

Now, we write X = PY, i.e.,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1/3 & -2/7 \\ 0 & 1 & 1/7 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

we obtain

$$\left.\begin{aligned} x_1 &= y_1 + \frac{1}{3} y_2 - \frac{2}{7} y_3 \\ x_2 &= y_1 + \frac{1}{7} y_3 \\ x_3 &= y_2 \end{aligned}\right\} \quad \ldots (1)$$

Thus the linear transformation (1) transforms the quadratic form corresponding to the given symmetric matrix to the form

$$X'AX = 6x_1^2 + 3x_2^2 + 3x_3^2 - 4x_1x_2 - 2x_2x_3 + 4x_3x_1$$

to the diagonal form

$$Y'P'APY = 6y_1^2 + \frac{7}{3} y_2^2 + \frac{16}{7} y_3^2$$

Example 7.7 : If A is a symmetric and S a skew-symmetric matrix of order n such that A + S is non-singular, and

$$T = (A + S)^{-1} (A - S)$$

prove that $T' (A + S) T = A + S$, $T' (A - S) T = A - S$

Solution : We have, $T' = (A - S)' [(A + S)^{-1}]'$

$$= (A' - S') [(A + S)']^{-1} = (A + S) (A - S)^{-1}$$

$\Rightarrow$ $T' (A + S) T = (A + SD) (A - S)^{-1} (A + S) (A + S)^{-1} (A - S)$

$$= A + S \quad \ldots (i)$$

Similarly, $T' (A - S) T = A - S \quad \ldots (ii)$

Hence the result.

Example 7.8 : Show that a non-singular symmetric matrix A is congruent to its inverse.

Solution : Let A be a non-singular symmetric matrix so that

$$A' = A \text{ and } A^{-1} \text{ exists}$$

To prove that A is congruent to its inverse A^{-1}, we must show that $\exists$ non-singular matrix P such that

$$A = P'A^{-1}P$$

write $P = B'B$ with $BAB' = I \quad \ldots (1)$

This proves that P is non-singular and symmetric.

$$P'AP = PAB \qquad [\because P' = P]$$

$$= B'BAB'B = B' (BAB') B$$

$$= B'IB = B'B = P$$

Finally, $P'AP = P$ … (2)

$(P'AP)\,P^{-1} = PP^{-1}$

$(P'A)\,(PP^{-1}) = I$

or $P'A = I$

or $P' = IA^{-1} = A^{-1}$

or $P' = A^{-1}$ … (3)

$P'AP = PAP = A^{-1}AA^{-1} = IA^{-1} = A^{-1}$

Thus, $P'AP = A^{-1}$

or $A = (P')^{-1}\,A^{-1}\,P^{-1}$

$= (P^{-1})'\,A^{-1}\,P^{-1}$

$= Q'A^{-1}\,Q$ on taking P^{-1}

Thus, $A = Q'A^{-1}$, Q is non-singular.

Hence, the required result forms.

Exercise 7.5

1. Reduce into a diagonal form by congruent elementary transformations to the following symmetric matrices :

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \\ -1 & 3 & 1 \end{bmatrix},\ B = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 2 & 3 \\ -1 & 3 & 1 \end{bmatrix},\ C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 0 & 1 \\ 3 & 0 & 1 & -2 \\ 4 & 1 & -2 & 3 \end{bmatrix}$$

2. Express the following quadratic form into a symmetric matrix and reduce into diagonal form by using elementary congruent transformation.
   (i) $10x^2 + y^2 + z^2 - 6xy - 2yz + 6xz$.
   (ii) $2x^2 + 9y^2 + 6z^2 + 8xy + 8yz + 6zx$.
   (iii) $3x^2 + 2y^2 + z^2 + 4xy + 2yz - 2xt + 2xz - 2zt$.
   (iv) $x_1^2 + 2x_2 - 3x_3^2 + 4x_4^2 + 2x_1x_2 - 4x_3x_2 - 2x_4x_2 + 4x_3x_4$.
   (v) $3x_1^2 + 5x_2^2 - x_3^2 - 14x_4^2 - 8x_1x_2 + 2x_1x_3 - 10x_1x_4 + 2x_2x_4$.

7.8 Congruence Reduction of Skew-Symmetric Matrices

Thoerem 6 : If A be any skew-symmetric matrix over F, then there exists a non-singular matrix P over F such that

$$P'AP = \text{Diag. } [J, J, \ldots, J, 0, 0, \ldots, 0]$$

where, $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

The zeros in P'AP may or may not be present.

Firstly, we notice that every matrix congruent to a skew-symmetric matrix is itself skew, for

$$A' = -A \Rightarrow (P'AP)' = P'A'P = -(P'AP)$$

In particular, every matrix obtained by subjecting a skew-symmetric matrix to any elementary congruent transformation is skew.

We know that every diagonal element of a skew matrix is zero.

Suppose that $A \neq 0$, so that there exists some non-zero element a_{ij} of A.

Performing the pair of elementary congruent transformations

$$R_{i1}, C_{i1}, R_{j2}, C_{j2}$$

we get a skew matrix B congruent to A such that

$$b_{12} = a_{ij} \neq 0, \; b_{21} = -b_{12}$$

Performing the elementary congruent transformations

$$R_1 (-b_{12}^{-1}) = R_1 (b_{21}^{-1}), \; C_1 (b_{21}^{-1})$$

we get a skew matrix C congruent to B such that

$$C_{12} = -1, \; C_{21} = I$$

Performing now the series of pairs of elementary congruent transformations

$$R_{i2} (-C_{i1}) = R_{i2} (c_{1i}), C_{i2} (-c_{1i})$$

$$R_{i1} (c_{i1}) = R_{1i} (-c_{1i}), C_{i1} (-c_{1i})$$

i varying from 3 to n, we get a skew-symmetric matrix D such that

$$D = \begin{bmatrix} J & 0 \\ 0 & A_1 \end{bmatrix} = \text{Diag. } [J, A_1]$$

where, A_1 is an $(n-2)$ rowed skew-symmetric matrix.

If, now, $A_1 \neq 0$, we can deal with A_1 as we did with A without affecting the first two rows and columns of D.

Thus proceeding, we shall obtain a skew-symmetric matrix of the given form congruent to the given matrix A.

As at each stage of the above process, we have employed only elementary congruent transformations, we see at once the truth of the result as stated.

Corollary 1 **:** The rank of every skew-symmetric matrix is even.

This follows from the fact that the rank of P'AP is twice the number of times the matrix. J, appears.

Corollary 2 **:** Two n-rowed skew-symmetric matrixes over an arbitrary field F are congruent if, and only if, they have the same rank.

This is an immediate consequence of the theorem and the fact that the relation of congruence possesses symmetry and transitivity.

Note : Every n-rowed skew-symmetric matrix of the form

$$\text{Diag. } (J, J, \ldots, J, 0, \ldots 0)$$

is a canonical matrix for congruence, for every n-rowed skew-symmetric matrix is congruent to one and only one such canonical matrix.

The total number of n-rowed skew-symmetric canonical matrices is $\frac{1}{2}(n + 1)$ or $\frac{1}{2}(n + 2)$ according to n is odd or even.

Exercise 7.4

1. Reduce the canonical forms the following skew-symmetric matrices :

(i) $\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & 3 \\ 2 & -3 & 0 \end{bmatrix}$, (ii) $\begin{bmatrix} 0 & 1 & 4 \\ -1 & 0 & 2 \\ -3 & -2 & 0 \end{bmatrix}$, (iii) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ -1 & 0 & 4 & 5 \\ -2 & -4 & 0 & 6 \\ -3 & -5 & -6 & 0 \end{bmatrix}$.

2. Show that the determinant of every real skew-symmetric matrix is positive or zero.
3. Show that every skew-symmetric matrix of odd order is singular.

7.9 Properties of Definite, Semi-Definite and Indefinite Forms

7.9.1 Definite Forms

Let X'AX be a definite quadratic form.

Theorem : The form X'AX has the value 0, if and only if, X = 0. Also for a non-zero form the value is positive or negative according as it is positive or a negative form.

Let X'AX be a positive definite form in n variables, so that there exists a real non-singular linear transformations

$$X = PY$$

such that $X'AX = Y'P'APY = y_1^2 + \ldots + y_n^2$

Thus the form assumes only positive values except for those values of X which correspond to Y = 0 (i.e., $y_1 = 0$, $y_2 = 0$, …, $y_n = 0$) when it assumes the value 0. Also clearly, 0, is the only value of X corresponding to Y = 0.

The case of a negative definite form may be disposed of in a similar manner.

7.9.2 Semi-definite Forms

Theorem : Every value of a positive semi-definite form $X'AX \geq 0$. Also there exist non-zero values of X for which of the form assumes the value 0.

Let X'AX be positive semi-definite so that there exists a real non-singular transformation X = PY such that

$$X'AX = Y'P'APY = y_1^2 + \ldots + y_1^2, \ (r < n)$$

Thus the form assumes only positive values except for those values of X which correspond to the values of Y with first r components $y_1, \ldots, y_r$ all zero, when it assumes the value zero.

The linear transformation X = PY being non-singular, Y = 0 is the only value of Y which corresponds to X = 0. Thus every non-zero Y whose first r components are zero determines non-zero X and for every such non-zero value of X, the form X'AX becomes zero.

We may discuss the case of negative semi-definite forms in a similar manner and it may be shown that every value of a negative semi-definite form $X'AX \leq 0$ and that there exist non-zero X for which the form has the value 0.

7.9.3 Indefinite Forms

Theorem : An indefinite form assumes positive as well as negative values.

For an indefinite form X'AX, there exists a real non-singular linear transformation X = PY such that

$$X'AX = Y'P'APY = y_1^2 + \ldots + y_s^2 - y_{s+1}^2 - \ldots - y_r^2 \ (s - r \leq n)$$

The value of X which corresponds to non-zero value of Y for which

$$y_1 = 0, \ldots, y_s = 0; \ y_{r+1} = 0, \ldots, y_n = 0$$

makes the form negative and the value of X which corresponds to a non-zero value of Y for which

$$y_{s+1} = 0, \ldots, y_r = 0; \ y_{r+1} = 0, \ldots, y_n = 0$$

makes the form positive.

SOLVED EXAMPLES

Example 7.9 : Show that a real symmetric matrix is positive definite if all its eigen values are positive.

Solution : Let A be a real symmetric matrix of order n. Then there exists an orthogonal matrix P such that

$$P^{-1}AP = P'AP = D = \text{Diag.} \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$$

where, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigen values of A.

Let X'AX be the real quadratic form corresponding to the matrix A. Let us transform this quadratic form by the real non-singular linear transformation X = PY where, $Y = [y_1, y_2, \ldots, y_n]'$. Then

$$X'AX = (PY)' A(PY) = Y'P'APY = Y'P$$

$$\therefore \quad X'AX = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \quad \ldots (1)$$

Now suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive. Then the right hand side of (1) ensures that $X'AX \geq 0$ for all real vector X. Also

$\Rightarrow \quad \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 = 0$

$\Rightarrow \quad y_1 = y_2 = \ldots = y_n = 0 \quad [\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive$]$

$\Rightarrow \quad Y = 0$

$\Rightarrow \quad P^{-1}X = 0 \quad [\because X = PY \Rightarrow Y = P^{-1} X]$

$\Rightarrow \quad P(P^{-1}X) = P0$

$\Rightarrow \quad X = 0$

Thus if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive, then X'AX is positive definite and so that matrix X is positive definite.

Conversely suppose that A is a positive definite matrix. Then the quadratic form X'AX is positive definite. Hence, $X'AX \geq 0$ for all real vectors X.

$\Rightarrow \quad \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \geq 0$ for all real vectors.

$\Rightarrow \quad \lambda_1, \lambda_2, \ldots, \lambda_n$ are all ≥ 0.

Also X'AX = 0 only if X = 0.

$\Rightarrow \quad \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 = 0$ only if PY = 0.

$\Rightarrow \quad \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 = 0$ only if Y = 0.

$\Rightarrow \quad \lambda_1, \lambda_2, \ldots, \lambda_n$ are all not equal to zero.

Example 7.10 : Positive definite real symmetric matrix is non-singular.

Solution : Suppose A is a positive definite real symmetric matrix. Then the eigen values of A are all positive. Also there exists an orthogonal matrix such that

$$P^{-1}AP = D$$

where, D is a diagonal matrix having the eigen values of A as its diagonal elements. So all diagonal elements of D are positive and thus D is non-singular.

Now, $A = PDP^{-1} \Rightarrow$ A is non-singular

Example 7.11 : Show that a real symmetric matrix A is positive definite if and only if there exists a non-singular matrix Q such that A = Q'A.

Solution : Suppose A is positive definite. Then all the eigen values of A are positive and we find an orthogonal matrix P such that

$$P^{-1}AP = P'AP = D = \text{Diag. } [\lambda_1, \lambda_2, \ldots, \lambda_n]$$

where, each $\lambda_i > 0$. Let $D_1 = \text{Diag. } [\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}]$. Then $D_1^2 = D$ and $D_i' = D_1$. We have

$$\begin{aligned} A &= PDP^{-1} = PD_1^2P^{-1} = PD_1D_1P' \\ &= (PD_1)\,(PD_1)' = Q'Q \text{ where, } Q = (PD_1)' \end{aligned}$$

obviously, Q is non-singular since P and D_1 are non-singular.

Conversely, suppose that A = Q'Q where, Q is non-singular. We have for all real vectors X,

$$\begin{aligned} X'AX &= X'Q'QX = (QX)'\,(QX) \\ &= Y'Y \end{aligned}$$

where, $Y = QX$ is a real n-vector

$$\geq 0$$

Also, $X'AX = 0 \Rightarrow Y'Y = 0$

$$\Rightarrow Y = 0$$

$$\Rightarrow QX = 0 \qquad [\because Y = QX]$$

$$\Rightarrow Q^{-1}\,(QX) = Q^{-1}Q$$

$$\Rightarrow X = 0$$

$\therefore$ X'AX is a positive definite real quadratic form and so the symmetric matrix A is positive definite.

Exercise 7.5

1. Show that the determinant of a semi-definite matrix is zero and that of a positive definite matrix is positive.
2. If A is a real symmetric singular matrix show that there exist non-zero X such that

   $$X'AX = 0$$

   [**Hint :** Every solution of AX = 0 makes X'AX = 0]
3. A is a real symmetric matrix such that |A| < 0, show that there exists a real column matrix X, such that

   $$X'AX < 0$$
4. If A is a real symmetric n-rowed matrix of rank r, show that the quadratic form X'AX can be expressed as a quadratic form in r varibles.
5. If B is a real non-singular matrix, show that B'B is a positive definite, matrix. If B is singular, then B'B is positive semi-definite.

7.9.4 Necessary and Sufficient Conditions for a Definite Form

Theorem : A necessary and sufficient condition for a real quadratic form X'AX to be positive definite is that the leading principal minors of the matrix A of the form are all positive.

Result : The determinant of the matrix of a positive definite form is positive.

If X'AX be positive definite, then there exists a real-singular matrix P such that

$$P'AP = I$$

$$\Rightarrow \quad |P'AP| = |I|$$

$$\Rightarrow \quad |P'|\,|A|\,|P| = I$$

$$\Rightarrow \quad |A| = 1/|P|^2, \text{ for } |P| = |P'| \neq 0$$

Thus |A| is positive.

We now attend to the main theorem.

The condition is necessary. Let the form X'AX be positive definite. Let, s, be a natural number less than or equal to r.

Putting $x_{s+1} = 0, \ldots, x_n = 0$

in the positive definite form X'AX, we obtain a positive definite quadratic form in, s, variable $x_1, \ldots, x_s$, the determinant of whose matrix is the leading principal minor of order s of A. Thus, by the lemma, every leading principal minor of the matrix of a positive definite quadratic form is positive.

The condition is sufficient. We will employ the principle of mathematical induction for establishing the sufficient.

$$a_{11}x_1^2$$

is positive definite when a_{11} is positive.

Suppose, now, that the theorem is true for quadratic forms in m variables.

Consider any quadratic form in (m + 1) variables with corresponding symmetric matrix S such that the leading principal minors of its matrix, S, are all positive. We partition S as follows :

$$S = \begin{bmatrix} B & B_1 \\ B_1' & k \end{bmatrix}$$

where, B is an $m \times m$ symmetric matrix and B_1 is an $m \times 1$ column matrix.

The leading principal minors of B are all positive i.e., the leading principal minors of B and the determinant |S| are all positive.

As the theorem is assumed to be true for quadratic forms in m variables, there exist a non-singular matrix P of order m such that

$$P'BP = I_m$$

Let C, be an $m \times 1$ column matrix to be shortly determined in a suitable manner. We have

$$\begin{bmatrix} P' & 0 \\ C' & 1 \end{bmatrix}\begin{bmatrix} B & B_1 \\ B_1' & k \end{bmatrix}\begin{bmatrix} P & C \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} P'BP & P'BC + P'B_1 \\ C'BP + B_1'P & C'BC + C'B_1 + B_1' + k \end{bmatrix}$$

Consider $P'BC + P'B_1 = 0$

$\Leftrightarrow$ $BC + B_1 = 0$ for P' is non-singular

$\Leftrightarrow$ $C = -B^{-1}B_1$ for B is non-singular

With C determined as above, we have

$$P'BC + P'B_1 = 0,\ C'BP = B_1'P = 0,\quad C'BC = C'B_1 = 0$$

Thus, $\begin{bmatrix} P' & 0 \\ C' & 1 \end{bmatrix} S \begin{bmatrix} P & C \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} BP'P & 0 \\ 0 & B_1'C + k \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & B_1'C + k \end{bmatrix}$

Taking determinants, we obtain

$$|P'|\,|S|\,|P| = |I_m|\,|B_1'C + k| = B_1'C + k$$

for $B_1'C + k$ is a single element matrix.

$\therefore$ $|S|\,|P|^2 = B_1'C + k$

As $|S|$ is positive, and $|P| \neq 0$, we see that $B_1'C + k$ is also positive.

We write $B_1'C + k = \beta^2$

here β is real. Thus,

$$\begin{bmatrix} P' & 0 \\ C' & I \end{bmatrix}\begin{bmatrix} P & C \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & \beta^2 \end{bmatrix}$$

Pre-multiplying and post-multiplying with

$$\begin{bmatrix} I_m & 0 \\ 0 & \beta^{-1} \end{bmatrix}$$

We have, $\begin{bmatrix} I_m & 0 \\ 0 & \beta^{-1} \end{bmatrix}\begin{bmatrix} P' & 0 \\ C' & I \end{bmatrix} S \begin{bmatrix} P & C \\ 0 & I \end{bmatrix}\begin{bmatrix} I_m & 0 \\ 0 & \beta^{-1} \end{bmatrix} = I_{m+1}$

Writing $Q = \begin{bmatrix} P & C \\ 0 & I \end{bmatrix}\begin{bmatrix} I_m & 0 \\ 0 & \beta^{-1} \end{bmatrix}$

$$\Rightarrow \qquad Q' = \begin{bmatrix} I_m & 0 \\ 0 & \beta^{-1} \end{bmatrix} \begin{bmatrix} P' & 0 \\ C' & I \end{bmatrix}$$

be obtain $Q'SQ = I_{m+1}$

As S is congruent to I_{m+1}, we see that the corresponding quadratic form m(m + 1) variables is positive definite.

The result, now, follows by induction.

Corollary : A necessary and sufficient condition for a real quadratic $\lambda'AX$, to be negative definite is that the leading principal minors of the matrix A, beginning from that of the first order, are alternately negative and positive for a form X'AX is negative definite if and only if, (−A)X is positive definite.

SOLVED EXAMPLES

Example 7.12 : Show that the form $x_1^2 + 2x_2^2 + 3x_3^2 + 2x_2x_3 - 2x_3x_1 + 2x_1x_2$ is indefinite and find two sets of values of x_1, x_2, x_3 for which the form assume the positive and negative values.

Solution : The matrix of the form is

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 3 \end{bmatrix}$$

We write $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

and perform elementary congruent transformations to find a diagonal matrix congruent to A.

Performing successively the elementary congruent transformations

$$R_{21}(-1), C_{21}(-1); R_{31}(1), C_{31}(1), R_{32}(-2), C_{32}(-2)$$

we obtain $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix} A \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$

Thus the linear transformation

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$\Leftrightarrow \quad \left.\begin{aligned} x_1 &= y_1 - y_2 + 3y_3 \\ x_2 &= y_2 - 2y_3 \\ x_3 &= y_3 \end{aligned}\right\} \quad \ldots (1)$$

transforms the given form to the diagonal form

$$y_1^2 + y_2^2 - 2y_3^2$$

so that the rank is 3 and index is 2. The form is obviously indefinite.

It is at once seen that

$$y_1 = 0,\ y_2 = 0,\ y_3 = 1$$

makes the form negative and

$$y_1 = 0,\ y_2 = 1,\ y_3 = 0$$

makes the form positive.

Substituting these values in (1), we see that the sets of values

$$x_1 = 3,\ y_2 = -2,\ x_3 = 1;\ x_1 = -1,\ y_2 = 1,\ x_3 = 0$$

respectively make the form negative and positive.

Example 7.13 : Show that the form

$$5x_1^2 + 26x_2^2 + 10x_3^2 + 4x_2x_3 + 14x_3x_1 + 6x_1x_2$$

is positive semi-definite and find a non-zero set of values of x_1, x_2, x_3 which makes the form zero.

Solution : The matrix of the form is

$$A = \begin{bmatrix} 5 & 3 & 7 \\ 3 & 26 & 2 \\ 7 & 2 & 10 \end{bmatrix}$$

We write $\begin{bmatrix} 5 & 3 & 7 \\ 3 & 26 & 2 \\ 7 & 2 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

To avoid fractions, we first perform the elementary congruent transformations

$$R_2(5), C_2(5); R_3(5), C_2(5)$$

and obtain

$$\begin{bmatrix} 5 & 15 & 35 \\ 15 & 650 & 50 \\ 35 & 50 & 250 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Performing successively the elementary congruent transformations

$$R_{21}(-3), C_{21}(-3); R_{31}(-7), C_{31}(-7); R_{23}(11), C_{23}(11)$$

we obtain

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -80 & 5 & 55 \\ -7 & 0 & 5 \end{bmatrix} A \begin{bmatrix} 1 & -80 & -7 \\ 0 & 5 & 0 \\ 0 & 55 & 5 \end{bmatrix}$$

so that the linear transformation

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & -80 & -7 \\ 0 & 5 & 0 \\ 0 & 55 & 5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$\Leftrightarrow \quad \left.\begin{aligned} x_1 &= y_1 - 80y_2 - 7y_3 \\ x_2 &= 5y_2 \\ x_3 &= 55y_2 + 5y_3 \end{aligned}\right\}$$

transforms the given form to

$$5y_1^2 + 5y_3^2$$

Thus the form is positive semi-definite.

It is obvious that the set of values

$$y_1 = 0,\ y_2 = 1,\ y_3 = 0$$

to which corresponds the set of values

$$x_1 = -80,\ x_2 = 5,\ x_3 = 55$$

makes the given form zero.

Example 7.14 : Verify the conclusions of the above theorems of the following matrices :

(i) $B = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 3 & -3 \\ 5 & 3 & 3 \end{bmatrix}$ (ii) $B = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix}$

Solution : (i) The rank of B is 2. Also

$$BB' = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 3 & -3 \\ 5 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 5 \\ -1 & 3 & 3 \\ 3 & -3 & 3 \end{bmatrix} = \begin{bmatrix} 11 & -11 & 11 \\ -11 & 19 & 5 \\ 11 & 5 & 43 \end{bmatrix}$$

It may now be shown by performing elementary congruent transformations that the matrix BB' is congruent to the diagonal matrix

Diag. (11 8 0)

so that BB' is a positive semi-definite matrix of rank 2.

(ii) The rank of B is 3. Also

$$BB' = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 3 \\ 1 & 2 & 2 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 8 & 3 \\ 8 & 14 & 8 \\ 3 & 8 & 11 \end{bmatrix}$$

As the three leading principal minors of BB' are all positive the matrix B' is positive definite.

Example 7.15 : If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ prove that $X'A'AX \geq 0$ for all real X and find the set of values of X for which $X'AAX = 0$.

Show that AA' is the matrix of a positive definite quadratic form in two variables.

Solution : The columns of A form a linearly dependent system. In fact, we have

$$C_3 = 2C_2 - C_1$$

where, C_1, C_2, C_3 denote the three columns of A. Thus, by A'A is a positive semi-definite matrix.

This result may also be shown directly by computing A'A.

We have $X'A'AX = (AX)'(AX)$

so that $X'A'AX = 0 \Leftrightarrow AX = 0$

The matrix equation $AX = 0$

is equivalent to the system of equations

$$x_1 + 2x_2 + 3x_3 = 0$$
$$4x_1 + 5x_2 + 6x_3 = 0$$

The equations are satisfied by

$$X = K \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

where, k is any number.

Again AA' is the Gram matrix of the matrix A', viz.

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

where two columns are linearly independent. Thus AA' is a positive definite 2×2 matrix.

Exercise 7.6

1. Classify the following forms as definite, semi-definite and indefinite :
   (i) $-y^2 - 2yz - 2xy$.
   (ii) $2x^3 + 2y^2 + 3z^2 - 4yz - 4zx + 2xy$.
   (iii) $26x^2 + 20y^2 + 10z^2 - 4yz - 16zx - 36xy$.
   (iv) $x_1^2 + 4x_2^2 + x_1^3 - 4x_2x_3 + 2x_3x_1 - 4x_1x_2$.
   (v) $9x_1^2 + 4x_2^2 + 4x_3^2 + x_4^2 + 8x_2x_3 + 12x_3x_1 + 12x_1x_2$.
2. Show that the following forms are indefinite and find sets of values of the variables for which they assume positive, negative and zero values.
   (i) $11x^2 + 14xy + 8yz + 14xz$, (ii) $x^2 + y^2 + 7z^2 - 16xy\ 8yz - 8zx$.
3. Show that the following forms are semi-definite and find non-zero sets of values of the variables for which they assume zeros as their values :
   (i) $6x^2 + 17y^2 + 3z^2 - 20xy - 14yz + 8zx$.
   (ii) $5 - x^2 - 5y^2 - 14z^2 - 2xy = 16yz + 8zx$.
4. Prove that the quadratic form
   $6x_1^2 + 3x_2^2 + 14x_3^2 + 4x_2x_3 + 18x_3x_1 + 4x_1x_2$ is positive definite.

Important Points

- Definition of quadratic form, representation in matrix form.
- Positive definite quadratic form and positive definite matrix.
- Maximum and minimum values of quadratic form.
- A symmetric matrix A is positive definite if and only if all the eigenvalue of A are positive.
- A symmetric matrix A is positive definite if and only if the determinant of every principal submatrix is positive.

Miscellaneous Exercises

(A) Theory Question :

1. What is quadratic form ? State the nature of quadratic form.
2. How to reduce a symmetric matrix into a quadratic form ?
3. State the properties of definite matrices.
4. Define : (i) Quadratic form, (ii) Congruence of quadratic form.

(B) Numerical Problems :

1. Which of the following matrices are positive definite and which are negative definite

   (a) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 \\ 4 & 3 \end{bmatrix}$

2. Which of the following matrices are positive definite and which are negative definite

(a) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -1 & \frac{2}{3} & 0 \\ 5 & -8 & -\frac{1}{4} \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$

3. Find the symmetric matrix that corresponds to each of the following quadratic forms.

(a) $3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$.

(b) $3x^2 + xz - 2yz$

(c) $2x^2 - 5y^2 - 7z^2$.

Ans. (a) $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & -1 & -3 \\ 4 & -3 & 1 \end{bmatrix}$ (b) $B = \begin{bmatrix} 3 & 0 & 1/2 \\ 0 & 0 & -1 \\ 1/2 & -1 & 0 \end{bmatrix}$ (c) $C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -7 \end{bmatrix}$

4. Find the quadratic form that corresponds to each of the following symmetric matrices :

$A = \begin{bmatrix} 5 & -3 \\ -3 & 8 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -5 & 7 \\ -5 & -6 & 8 \\ 7 & 8 & -9 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 4 & -1 & 5 \\ 4 & -7 & -6 & 8 \\ -1 & -6 & 3 & 9 \\ 5 & 8 & 9 & 1 \end{bmatrix}$

5. Determine whether each of the following quadratic forms is positive definite

(a) $x^2 + 2y^2 - 4xz - 4yz + 7z^2$

(b) $x^2 + y^2 + 2xz + 4yz + 3z^2$

6. Determine whether or not each of the following quadratic forms is positive definite :

(a) $x^2 + 7y^2 - 4xy$

(b) $x^2 + 8xy + 5y^2$

(c) $3x^2 + 2xy + y^2$

7. Suppose A is real symmetric positive definite matrix. Show that A = P'P for some non-singular matrix P.

Answers

1. (a) Neither positive definite nor negative definite.
   (b) Negative definite.
   (c) Semi positive definite.
   (d) Negative definite
2. (a) Positive definite
   (b) Neither negative definite nor positive
   (c) Positive definite.

❐❐❐

APPENDIX

APPENDIX A : SOLVED PROBLEMS

(A) DETERMINANT AND RANK OF MATRIX

Q. 1 Compute the determinant and rank of each of the following matrices :

(a) $\begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 6 & 9 & 1 \end{bmatrix}$

Solution : $A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 6 & 9 & 1 \end{bmatrix}$

$= 2(6 \times 1 - 9 \times 7) - 3(5 \times 1 - 6 \times 7) + 4(5 \times 9 - 6 \times 6)$

$= -114 + 153 - 12$

$= 27$

$\therefore \quad |A| = 27$

To find the rank of matrix A we make row transformations.

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 6 & 9 & 1 \end{bmatrix} \xrightarrow{R_1/2} \begin{bmatrix} 1 & 3/2 & 2 \\ 5 & 6 & 7 \\ 8 & 9 & 1 \end{bmatrix}$$

$$\xrightarrow[R_3 = R_3 - 8R_1]{R_2 = R_2 - 5R_1} \begin{bmatrix} 1 & 3/2 & 2 \\ 0 & -3/2 & -3 \\ 0 & -3 & -15 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & 3/2 & 2 \\ 0 & -3/2 & -3 \\ 0 & 0 & -9 \end{bmatrix}$$

$$\xrightarrow[-\frac{R_3}{9}]{-\frac{2}{3}R_2} \begin{bmatrix} 1 & 3/2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \quad \therefore \text{ rank (A) = (3).}$$

(b) $\begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ 2 & 4 & 2 & 4 & 2 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 3 & 1 \end{bmatrix}$

Solution : To find the rank of matrix B we make row transformations.

$$\xrightarrow{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ 0 & -2 & -1 & -10 & -16 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 3 & 1 \end{bmatrix}$$

$$R_2 \to -\frac{R_2}{2} \quad \begin{bmatrix} 1 & 3 & 5 & 7 & 9 \\ 0 & 1 & 1/2 & 5 & 8 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 3 & 1 \end{bmatrix}$$

Rank (B) = 4

Determinant :

$$B = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 2 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 3 \end{bmatrix}$$

$$|B| = 1\begin{bmatrix} 4 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix} - 2\begin{bmatrix} 3 & 5 & 7 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix} + 0\begin{bmatrix} 3 & 5 & 7 \\ 4 & 2 & 4 \\ 0 & 2 & 3 \end{bmatrix} - 0\begin{bmatrix} 3 & 5 & 7 \\ 4 & 2 & 4 \\ 0 & 1 & 2 \end{bmatrix}$$

$$= 1\,[4\,(3-4) - 2\,(0-0) + 4\,(0-0)] - 2\,[3\,(3-4) - 0\,(15-14) + 0\,(10-7)]$$

$$= 1\,[-4] - 2\,[-3]$$

$$= -4 + 6$$

$$\therefore \quad |B| = 2$$

Q. 2 Compute the determinants of each of the following matrices by reduction order method also compute the rank of matrix.

(i) $$A = \begin{bmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & 6 & 4 & 3 \end{bmatrix}$$

Solution : Use row reduction to obtain |A|.

$$|A| = \begin{bmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 2 \\ -1 & -6 & 4 & 3 \end{bmatrix}$$

$$\xrightarrow[R_2 = R_2 + 2R_3]{R_1 = R_1 - 2R_3} \begin{bmatrix} 0 & -1 & 1 & -6 \\ 0 & 3 & -2 & -1 \\ 1 & 3 & -2 & 2 \\ 0 & -3 & 2 & 5 \end{bmatrix}$$

Use $a_{31} = 1$ as a pivot 0's in first column.

$$|A| = \begin{vmatrix} 1 & 1 & -6 \\ 3 & -2 & -1 \\ -3 & 2 & 5 \end{vmatrix}$$

$$= -[-10 + 2] - 1[15 - 3] - 6[6 - 6] = +8 - 12 = -4$$

$\therefore \quad |A| = -4$

To compute the rank we use row reduction method.

$$\xrightarrow[R_2 = R_2 + 2R_3]{R_1 = \frac{R_1}{2}} \begin{bmatrix} 1 & 5/2 & -3/2 & -1 \\ 0 & 3 & -2 & -1 \\ 1 & 3 & -2 & 2 \\ -1 & 6 & 4 & 3 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 + R_1]{\substack{R_2 = R_2/3 \\ R_3 = R_3 + R_4}} \begin{bmatrix} 1 & 5/2 & -3/2 & -1 \\ 0 & 1 & -2/3 & -1/3 \\ 0 & 9 & 2 & 5 \\ 0 & 17/2 & 5/2 & 2 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 + \frac{17}{2} R_2]{R_3 = R_3 + 3R_2} \begin{bmatrix} 1 & 5/2 & -3/2 & -1 \\ 0 & 1 & -2/3 & -1/3 \\ 0 & 12 & 0 & 4 \\ 0 & 12 & 29/6 & 5/6 \end{bmatrix}$$

$$\xrightarrow{R_3 \leftrightarrow R_4} \begin{bmatrix} 1 & 5/2 & -3/2 & -1 \\ 0 & 1 & -2/3 & -1/3 \\ 0 & 12 & 29/6 & 5/6 \\ 0 & 12 & 0 & 4 \end{bmatrix}$$

$$\xrightarrow[R_4 \to \frac{R_4}{4}]{R_3 \to R_3 - R_4} \begin{bmatrix} 1 & 5/2 & -3/2 & -1 \\ 0 & 1 & -2/3 & -1/3 \\ 0 & 0 & 29/6 & -19.6 \\ 0 & 3 & 0 & 1 \end{bmatrix}$$

rank (A) = 4

(ii) $$B = \begin{bmatrix} 6 & 2 & 1 & 0 & 5 \\ 3 & 1 & 1 & -2 & 1 \\ 1 & 1 & 2 & -2 & 3 \\ 3 & 0 & 2 & 3 & -1 \\ -1 & -1 & -3 & 4 & 2 \end{bmatrix}$$

Solution : Use row reduction to obtain |B| of given matrix B.

$$\xrightarrow[R_3 = R_3 - R_2 \\ R_5 = R_5 + R_2]{R_1 = R_1 - 2R_2} \begin{bmatrix} 0 & 0 & -1 & 4 & 3 \\ 3 & 1 & 1 & 2 & 1 \\ -2 & 0 & 1 & 0 & 2 \\ 3 & 0 & 2 & 3 & -1 \\ 2 & 0 & -2 & 2 & 3 \end{bmatrix}$$

$$\therefore \begin{bmatrix} 0 & -1 & 4 & 3 \\ -2 & 1 & 0 & 2 \\ 3 & 2 & 3 & -1 \\ 2 & -2 & 2 & 3 \end{bmatrix}$$

Expanding above matrix along 1st column :

$$|B| = a_{11} C_{11} + a_{21} C_{21} + a_{31} C_{31} + a_{41} \cdot C_{41}$$
$$= 0\, C_{11} + (-2)\, C_{21} + (3)\, C_{31} + (2)\, C_{41}$$

$$C_{21} = (-1)^{2+1} \begin{vmatrix} -1 & 4 & 3 \\ 2 & 3 & -1 \\ -2 & 2 & 3 \end{vmatrix}$$

$$= (-1)\,[-1\,(9+2) - 4\,(6-2) + 3\,(4+6)$$
$$= (-1)\,[-11 - 4\,(4) + 3\,(10)]$$
$$= (-1)\,[-11 - 16 + 30]$$

$$\therefore \quad C_{21} = -3$$

$$C_{31} = (-1)^{3+1} \begin{vmatrix} -1 & 4 & 3 \\ 1 & 0 & 2 \\ -2 & 2 & 3 \end{vmatrix}$$

$$= [-1\,(-4) - 4\,(3+4) + 3\,(2)]$$
$$= [4 - 4\,(-1) + 6]$$
$$= 4 - 28 + 6 = -18$$

$\therefore \quad C_{31} = -18$

$$C_{41} = (-1)^{4+1}\begin{vmatrix} -1 & 4 & 3 \\ 1 & 0 & 2 \\ 2 & 3 & -1 \end{vmatrix}$$

$= (-1)[-1(-6) - 4(-1-4) + 3(3)$

$= (-1)[6 - 4(-5) + 9]$

$= (-1)[6 + 20 + 9] = -35$

$C_{41} = -35$

$\therefore \quad |B| = (0) + (-2)(-3) + (3)(-18) + (2)(-35)$

$|B| = -118$

To find rank of matrix B.

$$\xrightarrow{C_1 \to C_3} \begin{bmatrix} 1 & 2 & 6 & 0 & 3 \\ 1 & 1 & 3 & -2 & 1 \\ 2 & 1 & 1 & -2 & 3 \\ 2 & 0 & 3 & 3 & -1 \\ -3 & -1 & -1 & 4 & 2 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 - 2R_1,\ R_5 = R_5 + 3R_1]{R_2 = R_2 - R_1,\ R_3 = R_3 - 2R_1} \begin{bmatrix} 1 & 2 & 6 & 0 & 5 \\ 0 & -1 & -3 & -2 & -4 \\ 0 & -3 & -11 & -2 & -7 \\ 0 & -4 & 1 & 3 & -11 \\ 0 & 5 & 17 & 4 & 17 \end{bmatrix}$$

$$\xrightarrow{-R_2} \begin{bmatrix} 1 & 2 & 6 & 0 & 5 \\ 0 & 1 & 3 & 2 & 4 \\ 0 & -3 & -11 & -2 & -7 \\ 0 & -4 & 1 & 3 & -11 \\ 0 & 5 & 17 & 4 & 17 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 + 4R_2,\ R_5 = R_5 - 5R_2]{R_3 = R_3 + 3R_2} \begin{bmatrix} 1 & 2 & 6 & 0 & 5 \\ 0 & 1 & 3 & 2 & 4 \\ 0 & 0 & -2 & 4 & 5 \\ 0 & 0 & 13 & 11 & 5 \\ 0 & 0 & 2 & -6 & 3 \end{bmatrix}$$

$$R_3 = \frac{-1}{2} R_3 \longrightarrow \begin{bmatrix} 1 & 2 & 6 & 0 & 5 \\ 0 & 1 & 3 & 2 & 4 \\ 0 & 0 & 1 & -2 & -5/2 \\ 0 & 0 & 13 & 11 & 5 \\ 0 & 0 & 2 & -6 & -3 \end{bmatrix}$$

$\therefore$ rank (B) = 5

Q. 3. Let $A = \begin{bmatrix} 2 & 1 & -3 & 4 \\ 5 & -4 & 7 & -2 \\ 4 & 0 & 6 & -3 \\ 3 & -2 & 5 & 2 \end{bmatrix}$

Find minor and sign minor of the sub-matrix

$$M = \begin{bmatrix} -4 & 7 \\ -2 & 5 \end{bmatrix}$$

Solution : Here, $M = \begin{bmatrix} -4 & 7 \\ -2 & 5 \end{bmatrix}$ is given

$\therefore$ The minor is $|M| = \begin{vmatrix} -4 & 7 \\ -2 & 5 \end{vmatrix} = -20 + 14 = -6$

$\therefore$ The signed minor is $(-1)^{2+4+2+3}$ |M| because the row subscript are 2 and 4 and the column subscripts are 2 and 3.

$\therefore$ Signed mionor

$(-1)^{2+4+2+3} |M| = -|M| = -(-6) = +6.$

Q. 4 Let $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$

(a) Find |A|, (b) adj A, (c) A^{-1} using adj A (d) showthat $|A| = |A^T|$.

Solution : $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$

(a) To find |A| :

$$|A| = 2[(-4 \times 5) - (-1 \times 2)] - 3[(0 \times 5) - (1 \times 2)] - 4[(0 \times -1) - (1 \times 4)$$

$$|A| = 2[-20 + 2] - 3[0 - 2] - 4[0 + 4]$$

$$|A| = 2[-18] - 3[-2] - 4[4]$$

$$|A| = -36 + 6 - 16$$

$$|A| = -46$$

(b) adj A :

$$M_{11} = + \begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} = -20 + 2 = -18 \qquad \therefore C_{11} = -18$$

$$M_{12} = - \begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} = 0 - 2 = -2 \qquad \therefore C_{12} = 2$$

$$M_{13} = + \begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} = 0 + 4 = 4 \qquad \therefore C_{13} = 4$$

$$M_{21} = - \begin{vmatrix} 3 & -4 \\ -1 & 5 \end{vmatrix} = 15 - 4 = 11 \qquad \therefore C_{21} = -11$$

$$M_{22} = + \begin{vmatrix} 2 & -4 \\ 1 & 5 \end{vmatrix} = 10 + 4 = 14 \qquad \therefore C_{22} = 14$$

$$M_{23} = - \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = -2 - 3 = -5 \qquad \therefore C_{23} = 5$$

$$M_{31} = + \begin{vmatrix} 3 & -4 \\ -4 & 2 \end{vmatrix} = 6 - 16 = -10 \qquad \therefore C_{31} = -10$$

$$M_{31} = - \begin{vmatrix} 2 & -4 \\ 0 & 2 \end{vmatrix} = 4 - 0 = 4 \qquad \therefore C_{32} = -4$$

$$M_{33} = + \begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8 - 0 = -8 \qquad \therefore C_{33} = -8$$

$$\text{Cofactor matrix} = \begin{bmatrix} -18 & 2 & 4 \\ -11 & 14 & 5 \\ -10 & -4 & 8 \end{bmatrix}$$

$$\therefore \quad \text{Adj A} = \begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & 8 \end{bmatrix}$$

(c) A^{-1} using adj A :

$$A^{-1} = \frac{1}{|A|} \text{adj A} = \frac{1}{-46} \begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & 8 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 9/23 & 11/46 & 5/23 \\ -1/23 & -7/23 & 2/23 \\ -2/23 & -5/46 & 4/23 \end{bmatrix}$$

(d) Show taht $|A| = |A^T|$:

$$A^T = \begin{bmatrix} 2 & 0 & 1 \\ 3 & -4 & -1 \\ -4 & 2 & 5 \end{bmatrix}$$

$$|A^T| = 2[(-4 \times 5) - (2 \times -1)] - 0 + 1[3 \times 2 - 16]$$

$$|A^T| = 2[-20 + 2] + 1[6 - 16]$$

$$|A^T| = 2[-18] + [-10]$$

$$|A^T| = -36 - 10$$

$$|A^T| = -46$$

$\therefore$ $|A| = -46$ and $|A^T| = -46$

$\therefore$ $|A| = |A^T| = 46$

Hence the proof.

---

5. Solve the system of equations using determinants :

$$2x - 3y - z = 1$$
$$3x + 5y + 2z = 8$$
$$x - 2y - 3z = 1$$

Solution : $$D = \begin{vmatrix} 2 & 3 & -1 \\ 3 & 5 & 2 \\ 1 & -2 & -3 \end{vmatrix}$$

$$D = 2[-15 + 4] - 3[-9 - 2] + (-1)[-6 - 5]$$

$$D = (2 \times -11) - 3 \times -11 - 1 \times -11$$

$$D = -22 + 33 + 11$$

$$D = 22$$

$\because$ $D \neq 0$ $\therefore$ The system has unique solution.

$$\begin{bmatrix} 2 & 3 & -1 \\ 3 & 5 & 2 \\ 1 & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ -1 \end{bmatrix}$$

The compute N_x, N_y and N_z we replace their coefficients of x, y, z in the matrix of coefficients by constant terms.

$$\therefore \quad D_x = \begin{vmatrix} 1 & 3 & -1 \\ 8 & 5 & 2 \\ -1 & -2 & -3 \end{vmatrix}$$

$$= 2(-15+4) - 3(-24-2) - 1(-16-5)]$$
$$= 2(-11) - 3(-26) - (-21)$$
$$= 88$$

$$D_y = \begin{vmatrix} 2 & 1 & -1 \\ 3 & 8 & 2 \\ 1 & 1 & -3 \end{vmatrix}$$

$$= 2(-24-2) - 1(-9-2) - 1(3-8)]$$
$$= 2(-26) - 1(-11) - 1(-5)]$$
$$= -36$$

$$D_z = \begin{vmatrix} 2 & 3 & 1 \\ 3 & 5 & 8 \\ 1 & -2 & 1 \end{vmatrix}$$

$$= 2(5+16) - 3(3-8) + -6-5)]$$
$$= 2(21) - 3(-5) + (-11)]$$
$$= 46$$

$$\therefore \quad x = \frac{D_x}{D} = \frac{88}{22} = 4$$

$$y = \frac{D_y}{D} = \frac{-36}{22} = -1.6363636$$

$$z = \frac{D_z}{D} = \frac{46}{22} = 2.090909$$

---

6. $B = \begin{bmatrix} 4 & -6 & 8 & 9 \\ 0 & -2 & 7 & -3 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

Using above matrix show that determinant of upper triangular matrix is the product of diagonal element.

Solution : The given matrix is $B = \begin{bmatrix} 4 & -6 & 8 & 9 \\ 0 & -2 & 7 & -3 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

To show that |B| = product of element.

First finding |B| by column expansion

$$|A| = a_{11} C_{11} + a_{21} C_{21} + a_{31} C_{31} + a_{41} C_{41}$$

$$|B| = 4 \cdot C_{11} + 0 + 0 + 0$$

$$\therefore \quad C_{11} = \begin{bmatrix} -2 & 7 & -3 \\ 0 & 5 & 6 \\ 0 & 0 & 3 \end{bmatrix} = [-2\,(15)] = -30$$

$$\therefore \quad |B| = 4 \times -30 = -120$$

$$|B| = 4 \times -2 \times 5 \times 3 = -120$$

Hence the proof.

---

7. Find the rank and basis of the row space of each of the following matrices :

(i) $A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{bmatrix}$

Solution : $A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 6 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{bmatrix}$

$\xrightarrow{R_2 - 2R_1}$ $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -3 \\ 3 & 10 & -6 & -5 \end{bmatrix}$ $\xrightarrow{R_3 = R_3 - 3R_1}$ $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 4 & -6 & -2 \end{bmatrix}$

$\xrightarrow{R_3 = R_3 - 2R_1}$ $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 2 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

The non-zero rows are the basis for row space.

In this case the basis is $\{(1, 2, 0, -1), (2, 6, 3, -3)\}$ and rank (A) = 2.

(ii) $B = \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{bmatrix}$

Solution : $B = \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{bmatrix}$

$$\xrightarrow[R_3 = R_3 - R_1]{R_2 = R_2 - R_1} \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 1 & 0 & -5 & -5 & 0 \\ 3 & 8 & 1 & -7 & 8 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 - 3R_1]{R_3 = R_3 - R_1} \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & -3 & -6 & -3 & 3 \\ 0 & -1 & -2 & -1 & 1 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 + R_2]{R_3 = R_3 + 3} \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & -3 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_3/-3} \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The non-zero rows are a basis form of the row space.

In this case the basis is {(1, 3, 1, –2, –3), (0, 1, 2, 1, –1)] and rank (B) = 2.

(iii) $$C = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 2 & 5 & 5 & 6 & 4 & 5 \\ 3 & 7 & 6 & 11 & 6 & 9 \\ 1 & 5 & 10 & 8 & 9 & 9 \\ 2 & 6 & 8 & 11 & 9 & 12 \end{bmatrix}$$

Solution : Let the given matrix is

$$C = \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 2 & 5 & 5 & 6 & 4 & 5 \\ 3 & 7 & 6 & 11 & 6 & 9 \\ 1 & 5 & 10 & 8 & 9 & 9 \\ 2 & 6 & 8 & 11 & 9 & 12 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 - R_2,\ R_5 = R_5 - R_2]{R_2 = R_2 - 2R_1,\ R_3 = R_3 - 6R_1} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ -3 & -5 & 0 & -7 & 0 & -3 \\ -1 & 0 & 5 & 2 & 5 & 4 \\ 0 & 1 & 3 & 5 & 5 & 7 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 - R_2,\ R_5 = R_5 - R_2]{R_3 = R_3 + 3R_1} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 1 & 3 & 2 & 3 & 3 \\ -1 & -1 & 2 & -1 & 3 & 3 \\ 0 & 0 & 0 & 8 & 3 & 6 \end{bmatrix}$$

$$\xrightarrow[R_4 = R_4 + R_1]{R_3 = R_3 - R_2} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 1 & 3 & 2 & 4 & 5 \\ 0 & 0 & 0 & 8 & 3 & 6 \end{bmatrix}$$

$$\xrightarrow{R_4 = R_4 - R_2} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 5 & 2 & 4 \\ 0 & 0 & 0 & 8 & 3 & 6 \end{bmatrix}$$

$$\xrightarrow[R_5 = R_5 - R_4]{R_4 = R_4 + 5R_3} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & -3 & -6 \\ 0 & 0 & 0 & 3 & 1 & 2 \end{bmatrix}$$

$$\xrightarrow[R_5 = R_5 - R_3]{R_4 = R_4/-3} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 4 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow[R_5 = R_5/4]{R_4 = R_4 + R_3} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow[R_5 = R_5 + R_4]{R_4 = R_4 + R_5} \begin{bmatrix} 1 & 2 & 1 & 3 & 1 & 2 \\ 0 & 1 & 3 & -3 & 2 & 1 \\ 0 & 0 & 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The non-zero rows are basis for the row space.

In this case the basis is {(1, 2, 1, 3, 1, 2), (2, 5, 5, 6, 4, 5), (3, 7, 6, 11, 6, 9)} and rank (C) = 3.

●●●

(B) INVERSE OF MATRIX

Find the inverse of each of the following matrices :

1. $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix}$.

To find the inverse of matrix A we proceed as follows :

$$AA^{-1} = I$$

$$\therefore \begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[(-1)\,R_2]{R_3 + 2R_1} \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & -4 \\ 0 & -2 & 7 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[R_3 + 2R_2]{R_1 + 2R_2} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & -4 \\ 0 & 0 & -1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & -1 & 0 \\ 2 & -2 & 1 \end{bmatrix}$$

$$\xrightarrow{(-1)\,R_3} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & -1 & 0 \\ -2 & 2 & -1 \end{bmatrix}$$

$$\xrightarrow{\substack{R_1 + 5R_3 \\ R_2 + 4R_3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -9 & 8 & -5 \\ -8 & 7 & -4 \\ -2 & 2 & -1 \end{bmatrix}$$

$$\therefore \quad A^{-1} = \begin{bmatrix} -9 & 8 & -5 \\ -8 & 7 & -4 \\ -2 & 2 & -1 \end{bmatrix}$$

---

2. $$A = \begin{bmatrix} 2 & -1 & 4 \\ -3 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

Solution : The given matrix is $\begin{bmatrix} 2 & -1 & 4 \\ -3 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}$.

To find the inverse of matrix A we proceed as follows :

$$A \cdot A^{-1} = I$$

$$\begin{bmatrix} 2 & -1 & 4 \\ -3 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{(-1)\,R_{31} \leftrightarrow R_1} \begin{bmatrix} 1 & -1 & 1 \\ -3 & 0 & 1 \\ 2 & -1 & 4 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\substack{R_2 + 3R_1 \\ R_3 - 2R_1}} \begin{bmatrix} 1 & -1 & 1 \\ 0 & -3 & 4 \\ 0 & 1 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & -3 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\xrightarrow{R_2 + 4R_3} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 12 \\ 0 & 1 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 4 & 1 & 5 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\xrightarrow{\substack{R_1 + R_2 \\ R_3 - R_2}} \begin{bmatrix} 1 & 0 & 13 \\ 0 & 1 & 12 \\ 0 & 0 & -10 \end{bmatrix} A^{-1} = \begin{bmatrix} 4 & 1 & 4 \\ 4 & 1 & 5 \\ -3 & -1 & -3 \end{bmatrix}$$

$$\xrightarrow{\left(-\frac{1}{10}\right)R_3} \begin{bmatrix} 1 & 0 & 13 \\ 0 & 1 & 12 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 4 & 1 & 4 \\ 4 & 1 & 5 \\ 3/10 & 1/10 & 3/10 \end{bmatrix}$$

$$\xrightarrow{\substack{R_2 - 12R_3 \\ R_1 - 13R_3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1/10 & -3/10 & 1/10 \\ 4/10 & -2/10 & 14/10 \\ 3/10 & 1/10 & 3/10 \end{bmatrix}$$

$$\therefore \quad A^{-1} = \begin{bmatrix} 1/10 & -3/10 & 1/10 \\ 4/10 & -2/10 & 14/10 \\ 3/10 & 1/10 & 3/10 \end{bmatrix}$$

---

3. $A = \begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$.

We have to find inverse of matrix A.

$$\therefore \quad AA^{-1} = I$$

$$\begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_3 \leftrightarrow R_1} \begin{bmatrix} 1 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\substack{R_2 - 2R_1 \\ R_3 - 2R_1}} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -2 & -3 \\ 0 & -3 & -2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -2 \\ 1 & 0 & -2 \end{bmatrix}$$

$$\xrightarrow{R_2 - R_3} \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & -1 \\ 0 & -3 & -2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

$$\begin{matrix} R_2 - 2R_2 \\ R_3 + 3R_2 \\ \longrightarrow \end{matrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & -5 \end{bmatrix} A^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 1 & 0 \\ -2 & 3 & -2 \end{bmatrix}$$

$$\begin{matrix} \left(-\frac{1}{5}\right) R_3 \\ \longrightarrow \end{matrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ -1 & 1 & 0 \\ 2/5 & 3/5 & 2/5 \end{bmatrix}$$

$$\begin{matrix} R_1 - 4R_3 \\ R_2 + R_3 \end{matrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 2/5 & 2/5 & -3/5 \\ -3/5 & 2/5 & 2/5 \\ 2/5 & -3/5 & 2/5 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 2/5 & 2/5 & -3/5 \\ -3/5 & 2/5 & 2/5 \\ 2/5 & -3/5 & 2/5 \end{bmatrix}$$

---

4. $A = \begin{bmatrix} 4 & -5 & 6 \\ -1 & 2 & 3 \\ -2 & 4 & 7 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 4 & -5 & 6 \\ -1 & 2 & 3 \\ -2 & 4 & 7 \end{bmatrix}$.

We have to find the inverse of matrix A.

$\therefore \quad AA^{-1} = I$

$$\begin{bmatrix} 4 & -5 & 6 \\ -1 & 2 & 3 \\ -2 & 4 & 7 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{matrix} (-1)\, R_2 \leftrightarrow R_1 \\ \longrightarrow \end{matrix} \begin{bmatrix} 1 & -2 & -3 \\ 4 & -5 & 6 \\ -2 & 4 & 7 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 4 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

$$\begin{matrix} \left(\frac{1}{3}\right) R_2 \\ \longrightarrow \end{matrix} \begin{bmatrix} 1 & -2 & -3 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1/3 & 4/3 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 + 2R_2} \begin{bmatrix} 1 & 0 & 9 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 2/3 & 5/3 & 0 \\ 1/3 & 4/3 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

$$\xrightarrow{\substack{R_2 - 6R_3 \\ R_1 - 9R_3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 2/3 & 59/3 & -9 \\ 1/3 & 40/3 & -6 \\ 0 & -2 & 1 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 2/3 & 59/3 & -9 \\ 1/3 & 40/3 & -6 \\ 0 & -2 & 1 \end{bmatrix}$$

---

5. $A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}$.

We have to find inverse of matrix A.

$\therefore \quad AA^{-1} = I$

$$\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{(-1)\, R_3 \leftrightarrow R_1} \begin{bmatrix} 1 & -1 & 1 \\ 2 & 3 & 2 \\ 2 & -2 & 4 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\substack{R_2 - 2R_1 \\ R_3 - 2R_1}} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\xrightarrow{R_2\left(\frac{1}{5}\right)} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1/5 & 2/5 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\xrightarrow{R_1 + R_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1/5 & -3/5 \\ 0 & 1/5 & 2/5 \\ 1 & 0 & 2 \end{bmatrix}$$

$$\xrightarrow{\left(\frac{1}{2}\right)R_3} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1/5 & -3/5 \\ 0 & 1/5 & 2/5 \\ 1/2 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 - R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -1/2 & 1/5 & -3/5 \\ 0 & 1/5 & 2/5 \\ 1/2 & 0 & 1 \end{bmatrix}$$

$$\therefore \quad A^{-1} = \begin{bmatrix} -1/2 & 1/5 & -8/5 \\ 0 & 1/5 & 2/5 \\ -1/2 & 0 & 1 \end{bmatrix}$$

---

6. $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \\ -4 & 5 & 2 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \\ -4 & 5 & 2 \end{bmatrix}$.

$\therefore \quad AA^{-1} = I$

$$\begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \\ -4 & 5 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_3 + 4R_1} \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \\ 0 & -3 & 14 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{(-1)\,R_2 - R_3} \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & -13 \\ 0 & -3 & 14 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & -1 & -1 \\ 4 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\substack{R_3 + 3R_2 \\ R_1 + 2R_2}} \begin{bmatrix} 1 & 0 & -23 \\ 0 & 1 & -13 \\ 0 & 0 & -25 \end{bmatrix} A^{-1} = \begin{bmatrix} -7 & -2 & -2 \\ -4 & -1 & -1 \\ -8 & -3 & -2 \end{bmatrix}$$

$$\xrightarrow{\left(-\frac{1}{25}\right)R_3} \begin{bmatrix} 1 & 0 & -23 \\ 0 & 1 & -13 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -7 & -2 & -2 \\ -4 & -1 & -1 \\ 8/25 & 3/25 & 2/25 \end{bmatrix}$$

$R_2 + 13R_3$, $R_1 + 23R_3$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 9/25 & 19/25 & -4/25 \\ 4/25 & 14/25 & 1/25 \\ 8/25 & 3/25 & 2/25 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 9/25 & 19/25 & -4/25 \\ 4/25 & 14/25 & 1/25 \\ 8/25 & 3/25 & 2/25 \end{bmatrix}$$

---

7. $$A = \begin{bmatrix} 1 & -2 & 3 & 4 \\ 3 & -1 & 2 & 5 \\ 2 & 4 & -5 & 1 \\ 4 & 2 & -1 & 3 \end{bmatrix}$$

Solution : The given matrix is $A = \begin{bmatrix} 1 & -2 & 3 & 4 \\ 3 & -1 & 2 & 5 \\ 2 & 4 & -5 & 1 \\ 4 & 2 & -1 & 3 \end{bmatrix}$.

We have to find inverse of matrix A.

$$A \cdot A^{-1} = I$$

$$\begin{bmatrix} 1 & -2 & 3 & 4 \\ 3 & -1 & 2 & 5 \\ 2 & 4 & -5 & 1 \\ 4 & 2 & -1 & 3 \end{bmatrix} A^{-} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$R_2 - 3R_1$, $R_3 - 2R_1$, $R_4 - 4R_1$ $\longrightarrow$

$$\begin{bmatrix} 1 & -2 & 3 & 4 \\ 0 & 5 & -7 & -7 \\ 0 & 8 & -11 & -7 \\ 0 & 10 & -13 & -13 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -4 & 0 & 0 & 1 \end{bmatrix}$$

$\left(\frac{1}{5}\right) R_2$ $\longrightarrow$

$$\begin{bmatrix} 1 & -2 & 3 & 4 \\ 0 & 1 & 7/5 & -7/5 \\ 0 & 8 & -11 & -7 \\ 0 & 10 & -13 & -13 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3/5 & 1/5 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -4 & 0 & 0 & 1 \end{bmatrix}$$

$R_1 + 2R_2$, $R_3 - 8R_2$, $R_4 - 10R_2$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & 1/5 & 6/5 \\ 0 & 1 & -7/5 & -7/5 \\ 0 & 0 & 1/5 & 21/5 \\ 0 & 0 & 1 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -1/5 & 2/5 & 0 & 0 \\ -3/5 & 1/5 & 0 & 0 \\ 14/5 & -8/5 & 1 & 0 \\ 2 & -2 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{5\,(R_3)} \begin{bmatrix} 1 & 0 & 1/5 & 6/5 \\ 0 & 1 & -7/5 & -7/5 \\ 0 & 0 & 1 & 21 \\ 0 & 0 & 1 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -1/5 & 7/5 & 0 & 0 \\ -3/5 & 1/5 & 0 & 0 \\ 14 & -8 & 5 & 0 \\ 2 & -2 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow[R_3 + 7/5\,R_3,\ R_4 - R_3]{R_1 - \left(\frac{1}{5}\right)R_3} \begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 28 \\ 0 & 0 & 1 & 21 \\ 0 & 0 & 0 & 20 \end{bmatrix} A^{-1} = \begin{bmatrix} -3 & 2 & -1 & 0 \\ 19 & 11 & 7 & 0 \\ 14 & -8 & 5 & 0 \\ -12 & 6 & -5 & 0 \end{bmatrix}$$

$$\xrightarrow{\left(-\frac{1}{20}\right)R_4} \begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 28 \\ 0 & 0 & 1 & 21 \\ 0 & 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -3 & 2 & -1 & 0 \\ 19 & -11 & 7 & 0 \\ 14 & -8 & 5 & 0 \\ -12/20 & 6/20 & -5/20 & 0 \end{bmatrix}$$

$$\xrightarrow[R_2 - 28R_4,\ R_3 - 21R_4]{R_1 + 3R_4} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} -96/20 & 58/20 & -35/20 & 0 \\ 716/20 & -388/20 & 280/20 & 0 \\ 532/20 & -286/20 & 205/20 & 0 \\ -12/20 & 6/20 & -5/20 & 0 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -96/20 & 58/20 & -35/20 & 0 \\ 716/20 & -388/20 & 280/20 & 0 \\ 532/20 & -286/20 & 205/20 & 0 \\ -12/20 & 6/20 & -5/20 & 0 \end{bmatrix}$$

---

8. $$A = \begin{bmatrix} 2 & -2 & 0 & -1 \\ 0 & 2 & 1 & 2 \\ 1 & -2 & 3 & -2 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

Solution : The given matrix is $A = \begin{bmatrix} 2 & -2 & 0 & -1 \\ 0 & 2 & 1 & 2 \\ 1 & -2 & 3 & -2 \\ 0 & 1 & 2 & 2 \end{bmatrix}$.

We have to find inverse of given matrix.

$\therefore \qquad A \cdot A^{-1} = I$

$$\begin{bmatrix} 2 & -2 & 0 & -1 \\ 0 & 2 & 1 & 2 \\ 1 & -2 & 3 & -2 \\ 0 & 1 & 2 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_3 \leftrightarrow R_1 \longrightarrow \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 2 & 1 & 2 \\ 2 & -2 & 0 & -1 \\ 0 & 1 & 2 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_3 - 2R_1 \longrightarrow \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 2 & 1 & 2 \\ 0 & 2 & -6 & 3 \\ 0 & 1 & 2 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_2 - R_4 \longrightarrow \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & 0 \\ 0 & 2 & -6 & 3 \\ 0 & 1 & 2 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{matrix} R_1 + 2R_2 \\ R_3 - 2R_2 \\ R_4 - R_2 \end{matrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -4 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 2 & 1 & -2 \\ 0 & 1 & 0 & -1 \\ 1 & -2 & -2 & 2 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$(-1)\, R_3 - R_4 \longrightarrow \begin{bmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 3 & 2 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 2 & 1 & -2 \\ 0 & 1 & 0 & -1 \\ -1 & 3 & 2 & -2 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\begin{matrix} R_1 - R_3 \\ R_2 + R_3 \\ R_4 - 3R_3 \end{matrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 17 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & -1 & -1 & 0 \\ -1 & 4 & 2 & -3 \\ -1 & 3 & 2 & -2 \\ 3 & -10 & -6 & 6 \end{bmatrix}$$

$$\left(\frac{1}{17}\right) R_4 \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & -1 & -1 & 0 \\ -1 & 4 & 2 & -3 \\ -1 & 3 & 2 & -2 \\ 3/17 & -10/17 & -6/17 & 6/17 \end{bmatrix}$$

$$\begin{matrix} R_1 - 3R_4 \\ R_2 + 5R_4 \\ R_3 + 5R_4 \\ \longrightarrow \end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 8/17 & 13/17 & 1/17 & -18/17 \\ -2/17 & 18/17 & 4/17 & -21/17 \\ -2/17 & 1/17 & 4/17 & -4/17 \\ 3/17 & -10/17 & -6/17 & 6/17 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 8/17 & 13/17 & 1/17 & -18/17 \\ -2/17 & 18/17 & 4/17 & -21/17 \\ -2/17 & 1/17 & 4/17 & -4/17 \\ 3/17 & -10/17 & -6/17 & 6/17 \end{bmatrix}$$

---

9. $A = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$. We have to find the inverse of given matrix.

$\therefore \quad A \cdot A^{-1} = I$

$$\begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_2 \leftrightarrow R_1 \quad \begin{bmatrix} 4 & -3 & 4 \\ 0 & 1 & -1 \\ 3 & -3 & 4 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_1 - R_3 \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 3 & -3 & 4 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R_3 - 3R_1 \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -3 & 4 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & -3 & 4 \end{bmatrix}$$

$$R_3 + 3R_2 \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 3 & -3 & 4 \end{bmatrix}$$

$R_2 + R_3$ $\quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$

$\therefore \quad A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$

---

10. $\quad A = \begin{bmatrix} 4 & 3 & 3 \\ -1 & 0 & 1 \\ -4 & -4 & -3 \end{bmatrix}$

Solution : The given matrix is $A = \begin{bmatrix} 4 & 3 & 3 \\ -1 & 0 & 1 \\ -4 & -4 & -3 \end{bmatrix}$. We have to find the inverse of given matrix.

$\therefore \quad A \cdot A^{-1} = I$

$$\begin{bmatrix} 4 & 3 & 3 \\ -1 & 0 & 1 \\ -4 & -4 & -3 \end{bmatrix} A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$(-1)\, R_2 \leftrightarrow R_1$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & -1 \\ 4 & 3 & 3 \\ -4 & -4 & -3 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$R_2 - 4R_1$, $R_3 + 4R_1$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 3 & 7 \\ 0 & -4 & -7 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 4 & 0 \\ 0 & -4 & 1 \end{bmatrix}$$

$(-1)\, R_2 - R_3$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & -4 & -7 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -4 & 1 \end{bmatrix}$$

$R_3 + 4R_2$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -7 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ -4 & -4 & -3 \end{bmatrix}$$

$$\xrightarrow{\left(-\frac{1}{7}\right)R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 4/7 & 4/7 & 3/7 \end{bmatrix}$$

$$\xrightarrow{R_1 + R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A^{-1} = \begin{bmatrix} 4/7 & -3/7 & 3/7 \\ -1 & 0 & -1 \\ 4/7 & 4/7 & 3/7 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 4/7 & -3/7 & 3/7 \\ -1 & 0 & -1 \\ 4/7 & 4/7 & 3/7 \end{bmatrix}$$

●●●

(C) SYSTEM OF LINEAR EQUATIONS

Q. Solve the following system of linear equations by using the following methods :

1. **Gauss-Elimination Method. OR**
2. **Gauss-Jordon Method.**

1. $x - 3y + z = -1$
$2x + y - 4z = -1$, and $\rho = 3$
$6x - 7y + 8z = 7$

Solution : The given system of equations in matrix form is as follows :

$$A = \begin{bmatrix} 1 & -3 & 1 \\ 2 & 1 & -4 \\ 6 & -7 & 8 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 \\ -1 \\ 7 \end{bmatrix}$$

$$\therefore \quad A|B = \left[\begin{array}{ccc|c} 1 & -3 & 1 & -1 \\ 2 & 1 & -4 & -1 \\ 6 & -7 & 8 & 7 \end{array}\right] \xrightarrow[R_3 - 6R_1]{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & -3 & 1 & -1 \\ 0 & 7 & -6 & 1 \\ 0 & 11 & 2 & 13 \end{array}\right]$$

$$\xrightarrow{\left(\frac{1}{7}\right)R_2} A|B = \left[\begin{array}{ccc|c} 1 & -3 & 1 & -1 \\ 0 & 1 & -6/7 & 1/7 \\ 0 & 11 & 2 & 13 \end{array}\right] \xrightarrow{R_3 - 11R_2} \left[\begin{array}{ccc|c} 1 & -3 & 1 & -1 \\ 0 & 1 & -6/7 & 1/7 \\ 0 & 0 & 80/7 & 80/7 \end{array}\right]$$

$$\rho(A|B) = \rho(A)$$

The system is consistent.

Now, $$|A| = \begin{bmatrix} 1 & -3 & 1 \\ 2 & 1 & -4 \\ 6 & -7 & 8 \end{bmatrix}$$

$$|A| = 1(8 - 28) + 3(16 + 24) + 1(-14 - 6)$$

$$|A| = 80 \neq 0$$

A^{-1} exist.

$\therefore$ By Gauss-Jordon method :

$$AX = B$$

$$\begin{bmatrix} 1 & -3 & 1 \\ 2 & 1 & -4 \\ 6 & -7 & 8 \end{bmatrix} X = \begin{bmatrix} -1 \\ -1 \\ 7 \end{bmatrix}$$

$R_2 - 2R_1$, $R_3 - 6R_1$ $\longrightarrow$

$$\begin{bmatrix} 1 & -3 & 1 \\ 0 & 7 & -6 \\ 0 & 11 & 2 \end{bmatrix} X = \begin{bmatrix} -1 \\ 1 \\ 13 \end{bmatrix}$$

$\left(\frac{1}{7}\right) R_2$ $\longrightarrow$

$$\begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & -6/7 \\ 0 & 11 & 2 \end{bmatrix} X = \begin{bmatrix} -1 \\ 1/7 \\ 13 \end{bmatrix}$$

$R_1 + 3R_2$, $R_3 - 11R_2$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & -11/7 \\ 0 & 1 & -6/7 \\ 0 & 0 & 80/7 \end{bmatrix} X = \begin{bmatrix} -4/7 \\ 1/7 \\ 80/7 \end{bmatrix}$$

$\left(\frac{7}{80}\right) R_3$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & -11/7 \\ 0 & 1 & -6/7 \\ 0 & 0 & 1 \end{bmatrix} X = \begin{bmatrix} -4/7 \\ 1/7 \\ 1 \end{bmatrix}$$

$R_1 + \frac{11}{7} R_3$, $R_2 + \frac{6}{7} R_3$ $\longrightarrow$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$X_1 = 1, \quad X_2 = 1, \quad X_3 = 1$$

---

2. $2x - 5y + 7z = 6$

$x - 3y + 4z = 3$, and $\rho = 2$

$3x - 8y + 11z = 11$

Solution : The system of given equations in the matrix form is as,

$$A = \begin{bmatrix} 2 & -5 & 7 \\ 1 & -3 & 4 \\ 3 & -8 & 11 \end{bmatrix} \text{ and } B = \begin{bmatrix} 6 \\ 3 \\ 11 \end{bmatrix}$$

$$A|B = \left[\begin{array}{ccc|c} 2 & -5 & 7 & 6 \\ 1 & -3 & 4 & 3 \\ 3 & -8 & 11 & 11 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & -3 & 4 & 3 \\ 2 & -5 & -7 & 6 \\ 3 & -8 & 11 & 11 \end{array}\right]$$

$$\xrightarrow[R_3 - 3R_1]{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & -3 & 4 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 2 \end{array}\right] \xrightarrow{R_3 - R_2} \left[\begin{array}{ccc|c} 1 & -3 & 4 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

$\rho(A) \neq \rho(A|B)$

$\therefore$ The system of equations is inconsistent. There is no solution.

---

3. $x + y + z = 7$

$x + 2y + 3z = 16$, and $\rho = 2$

$x - 2y + 3z = 5$

Solution : The given system of equations in matrix form is as follows :

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & -2 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 7 \\ 16 \\ 5 \end{bmatrix}$$

$$A|B = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 7 \\ 1 & 2 & 3 & 16 \\ 1 & 2 & 3 & 5 \end{array}\right]$$

$$\xrightarrow[R_3 - R_1]{R_2 - R_1} A|B = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 7 \\ 0 & 1 & 2 & 9 \\ 0 & 1 & 2 & -2 \end{array}\right]$$

$$\xrightarrow{R_3 - R_1} A|B = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 7 \\ 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & -11 \end{array}\right]$$

$\therefore \quad \rho(A) \neq \rho(A|B)$

$\therefore$ The system of equations is inconsistent.

---

4. $x + y + z = 2$

$x + 3y + 4z = 22$

$x + 3y + 6z = 11$, and $\rho = 3$

$x + 4y + 10z = 21$

Solution : The given system of equations in matrix form is as follows :

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \\ 1 & 3 & 6 \\ 1 & 4 & 10 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 \\ 22 \\ 11 \\ 21 \end{bmatrix}$$

$$A|B = \left[\begin{array}{cccc|c} 1 & 1 & 1 & & 2 \\ 1 & 3 & 4 & & 22 \\ 1 & 3 & 6 & & 11 \\ 1 & 4 & 10 & & 21 \end{array}\right] \begin{array}{l} R_2 - R_1 \\ R_3 - R_1 \\ \longrightarrow \\ R_4 - R_1 \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & & 2 \\ 0 & 2 & 3 & & 20 \\ 0 & 2 & 5 & & 9 \\ 0 & 3 & 9 & & 19 \end{array}\right]$$

$$\begin{array}{l} (-1)\, R_2 + R_4 \\ \longrightarrow \end{array} \left[\begin{array}{cccc|c} 1 & 1 & 1 & & 2 \\ 0 & 1 & 6 & & -1 \\ 0 & 2 & 5 & & 9 \\ 0 & 3 & 9 & & 19 \end{array}\right] \begin{array}{l} R_1 - R_2 \\ R_3 - 2R_2 \\ \longrightarrow \\ R_4 - 3R_2 \end{array} \left[\begin{array}{cccc|c} 1 & 0 & -5 & & 3 \\ 0 & 1 & 6 & & -1 \\ 0 & 0 & -7 & & 11 \\ 0 & 0 & -9 & & 22 \end{array}\right]$$

$$\begin{array}{l} R_4 - 2R_3 \\ \longrightarrow \end{array} A|B = \left[\begin{array}{cccc|c} 1 & 0 & -5 & & 3 \\ 0 & 1 & 6 & & -1 \\ 0 & 0 & -7 & & 11 \\ 0 & 0 & 5 & & 0 \end{array}\right] \begin{array}{l} \\ -\frac{1}{7} R_3 \\ \longrightarrow \\ \\ \end{array} \left[\begin{array}{cccc|c} 1 & 0 & -5 & & 3 \\ 0 & 1 & 6 & & -1 \\ 0 & 0 & 1 & & -11/7 \\ 0 & 0 & 5 & & 0 \end{array}\right]$$

$$\begin{array}{l} R_4 - 5R_3 \\ \longrightarrow \end{array} A|B = \left[\begin{array}{cccc|c} 1 & 0 & -5 & & 3 \\ 0 & 1 & 6 & & -1 \\ 0 & 0 & 1 & & -11/7 \\ 0 & 0 & 0 & & 55/7 \end{array}\right]$$

$\rho(A) \neq \rho(A|B)$

$\therefore$ The system of equations is inconsistent.

---

5. $3x + 4y - 6z + w = 7$

$x - 2y + 3z - 2w = -1$

$x - 3y + 4z - w = -2$

$5x - y + z - 2w = 4$

Solution : The given system of equations can be written in the matrix form as follows :

$$A = \left[\begin{array}{cccc} 3 & 4 & -6 & 1 \\ 1 & -2 & 3 & -2 \\ 1 & -3 & 4 & -1 \\ 5 & -1 & 1 & -2 \end{array}\right] \text{ and } B = \left[\begin{array}{c} 7 \\ -1 \\ -2 \\ 4 \end{array}\right]$$

$$A|B = \left[\begin{array}{cccc|c} 3 & 4 & -6 & 1 & 7 \\ 1 & -2 & 3 & -2 & -1 \\ 1 & -3 & 4 & -1 & -2 \\ 5 & -1 & 1 & -2 & 4 \end{array}\right]$$

$R_1 \leftrightarrow R_2$

$$A|B = \left[\begin{array}{cccc|c} 1 & -2 & 3 & -2 & -1 \\ 3 & 4 & -6 & 1 & 7 \\ 1 & -3 & 4 & -1 & -2 \\ 5 & -1 & 1 & -2 & 4 \end{array}\right] \xrightarrow[R_4 - 5R_1]{\substack{R_2 - 3R_1 \\ R_3 - R_1}} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -2 & -1 \\ 0 & 10 & -15 & 7 & 10 \\ 0 & -1 & 1 & 1 & -1 \\ 0 & 9 & -14 & 8 & 9 \end{array}\right]$$

$$\xrightarrow{(-1)\, R_3 \leftrightarrow R_2} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -2 & -1 \\ 0 & 1 & -1 & -1 & 1 \\ 0 & 10 & -15 & 7 & 10 \\ 0 & 9 & -14 & 8 & 9 \end{array}\right] \xrightarrow[R_4 - 9R_2]{R_3 - 10R_2} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -2 & -1 \\ 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & -5 & 17 & 0 \\ 0 & 0 & -5 & 17 & 0 \end{array}\right]$$

$$\xrightarrow{R_4 - R_3} A|B = \left[\begin{array}{cccc|c} 1 & -2 & 3 & -2 & -1 \\ 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & -5 & 17 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\rho(A) = \rho(A|B) = 3$$

The given system of equation is consistent.

By Gauss-elimination method.

$$AX = B$$

$$\left[\begin{array}{cccc} 3 & 4 & -6 & 1 \\ 1 & -2 & 3 & -2 \\ 1 & -3 & 4 & -1 \\ 5 & -1 & 1 & 2 \end{array}\right] X = \left[\begin{array}{c} 7 \\ -1 \\ -2 \\ 4 \end{array}\right]$$

$$\xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cccc} 1 & -2 & 3 & 2 \\ 3 & 4 & -6 & 1 \\ 1 & -3 & 4 & -1 \\ 5 & -1 & 1 & -2 \end{array}\right] X = \left[\begin{array}{c} -1 \\ 7 \\ -2 \\ 4 \end{array}\right]$$

$$\xrightarrow[R_4 - 5R_1]{\substack{R_2 - 3R_1 \\ R_3 - R_1}} \left[\begin{array}{cccc} 1 & -2 & 3 & -2 \\ 0 & 10 & -15 & 7 \\ 0 & -1 & 1 & 1 \\ 0 & 9 & -14 & 8 \end{array}\right] X = \left[\begin{array}{c} -1 \\ 10 \\ -1 \\ 9 \end{array}\right]$$

$$\xrightarrow{(-1)\,R_3 \leftrightarrow R_2} \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & 10 & -15 & 7 \\ 0 & 9 & -14 & 8 \end{bmatrix} X = \begin{bmatrix} -1 \\ 1 \\ 10 \\ 9 \end{bmatrix}$$

$$\xrightarrow{\substack{R_3 - 10R_2 \\ R_4 - 9R_2}} \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -5 & 17 \\ 0 & 0 & -5 & 17 \end{bmatrix} X = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\xrightarrow{R_4 - R_3} \begin{bmatrix} 1 & -2 & 3 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -5 & 17 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$x - 2y + 3z - 2w = -1$$
$$y - z - w = 1$$
$$-5z + 17w = 0$$
$$\therefore \quad 5z = 17w$$

Now take $w = t$

$$\therefore \quad z = \frac{17}{5}t$$
$$y - z - w = 1$$
$$y - \frac{17}{5}t - t = 1$$
$$y = 1 + \frac{22}{5}t$$
$$y = \frac{27}{5}t$$
$$x - 2y + 3z - 2w = -1$$
$$x - 2\frac{27}{5}t + 3\frac{17}{5}t - 2t = -1$$
$$x = -1 + \frac{13}{5}t$$
$$x = \frac{6}{5}t$$

The solution is $\begin{pmatrix} 6t/5 \\ 27t/5 \\ 17t/5 \\ t \end{pmatrix}$.

6. $4x + 3y - 7z + w = 7$
$2x + 5y - 7z - 3w = 3$ and $\rho = 4$
$8x - 3y + 5z - 2w = 0$
$6x + 3y + 9z - 4w = 9$

Solution : The given system of linear equations can be written in the matrix form is as follows :

$$A|B = \left[\begin{array}{cccc|c} 4 & 3 & -7 & 1 & 7 \\ 2 & 5 & -7 & -3 & 3 \\ 8 & -3 & 5 & -2 & 0 \\ 6 & 3 & 9 & -4 & 8 \end{array}\right]$$

$$\begin{matrix} R_1 + \frac{3}{10} R_4 \\ R_2 + \frac{23}{10} R_4 \\ R_3 + \frac{13}{10} R_4 \end{matrix} \rightarrow \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} x \\ y \\ z \\ w \end{array}\right] = \left[\begin{array}{c} 449/574 \\ 1037/574 \\ 251/574 \\ 434/287 \end{array}\right]$$

$\therefore \quad x = \frac{449}{574}, \; y = \frac{1037}{574}, \quad z = \frac{251}{574}$ and $w = \frac{434}{287}$

i.e. $x = 0.7822, \; y = 1.8066, \; z = 0.4372$ and $w = 1.5122$.

---

7. $x - 3y - 8z = -10$
$3x + y - 4z = 0,$ and $\rho = 2$
$2x + 5y + 6z = 13$

The system of equations in matrix form can be written as follows :

$$A|B = \left[\begin{array}{ccc|c} 1 & -3 & -8 & -10 \\ 3 & 1 & -4 & 0 \\ 2 & 5 & 6 & 13 \end{array}\right]$$

$$\begin{matrix} R_2 - 3R_1 \\ R_2 - 2R_1 \end{matrix} \longrightarrow \left[\begin{array}{ccc|c} 1 & -3 & -8 & -10 \\ 0 & 10 & 20 & 30 \\ 0 & 11 & 22 & 33 \end{array}\right] R_2\left(\frac{1}{10}\right) \left[\begin{array}{ccc|c} 1 & -3 & -8 & -10 \\ 0 & 1 & 2 & 3 \\ 0 & 11 & 22 & 33 \end{array}\right]$$

$R_3 - 11R_2$

$$\therefore \quad A|B = \left[\begin{array}{ccc|c} 1 & -3 & -8 & -10 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\rho(A) = \rho(A|B) = 2$

The system is consistent.

$\therefore$ By Gauss-elimination method

$$AX = B$$

$$\begin{bmatrix} 1 & -3 & -8 \\ 3 & 1 & -4 \\ 2 & 5 & 6 \end{bmatrix} X = \begin{bmatrix} -10 \\ 0 \\ 13 \end{bmatrix}$$

$$\xrightarrow[R_3 - 2R_1]{R_2 - 3R_1} \begin{bmatrix} 1 & -3 & -8 \\ 0 & 10 & 20 \\ 0 & 11 & 22 \end{bmatrix} X = \begin{bmatrix} -10 \\ 30 \\ 33 \end{bmatrix}$$

$$\xrightarrow{\left(\frac{1}{10}\right) R_2} \begin{bmatrix} 1 & -3 & -8 \\ 0 & 1 & 2 \\ 0 & 11 & 22 \end{bmatrix} X = \begin{bmatrix} -10 \\ 3 \\ 33 \end{bmatrix}$$

$$\xrightarrow{R_3 - 11R_3} \begin{bmatrix} 1 & -3 & -8 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -10 \\ 3 \\ 0 \end{bmatrix}$$

$$x - 3y - 8z = -10$$

$$y + 2z = 3$$

Take $z = s$

$$y + 2s = 3$$

$$y = 3 - 2s$$

$\therefore$ $x - 3y - 8s = -10$

$$x - 3(3 - 2s) - 8s = -10$$

$$x - 9 + 6s - 8s = -10$$

$$x = -10 + 9 + 2s$$

$$x = 2s - 1$$

The solution is $\begin{pmatrix} 2s - 1 \\ 3 - 2s \\ s \end{pmatrix}$

---

8. $x + y + z = 3$

$3x - 5y + 2z = 8$ and $\rho = 2$

$5x - 3y + 4z = 14$

In matrix form,

$$A|B = \left[\begin{array}{cccc|c} 1 & 1 & 1 & & 3 \\ 3 & -5 & 2 & & 8 \\ 5 & -3 & 4 & & 14 \end{array}\right]$$

$R_1 \leftrightarrow R_2$

$$A|B = \left[\begin{array}{cccc|c} 2 & 5 & -7 & 3 & 3 \\ 5 & 3 & -7 & 1 & 7 \\ 8 & -3 & 5 & -2 & 0 \\ 6 & 3 & 9 & -4 & 8 \end{array}\right]$$

$$\begin{matrix} R_2 - 2R_1 \\ R_3 - 4R_1 \\ \longrightarrow \\ R_4 - 3R_1 \end{matrix} \quad A|B = \left[\begin{array}{cccc|c} 2 & 5 & -7 & -3 & 3 \\ 0 & -7 & 7 & 7 & 1 \\ 0 & 17 & 33 & 10 & -12 \\ 0 & -12 & 30 & 5 & -1 \end{array}\right]$$

$$\begin{matrix} \left(-\frac{1}{7}\right) R_2 \\ \longrightarrow \end{matrix} \quad A|B = \left[\begin{array}{cccc|c} 2 & 5 & -7 & -3 & 3 \\ 0 & 1 & -1 & -1 & -1/7 \\ 0 & 17 & 33 & 10 & -12 \\ 0 & -12 & 30 & 5 & -1 \end{array}\right]$$

$$\begin{matrix} R_3 - 17R_2 \\ \longrightarrow \\ R_4 + 12R_2 \end{matrix} \quad A|B = \left[\begin{array}{cccc|c} 2 & 5 & -7 & -3 & 3 \\ 0 & 1 & -1 & -1 & -1/7 \\ 0 & 0 & 50 & 27 & -67/7 \\ 0 & 0 & 18 & 7 & -19/7 \end{array}\right]$$

$\therefore \quad \rho(A) = \rho(A|B)$

The system is consistent.

$\therefore$ The solution can be obtain by Gauss-Jordon methods follows :

$$AX = B$$

$$\left[\begin{array}{cccc} 4 & 3 & -7 & 1 \\ 2 & 5 & -7 & -3 \\ 8 & -3 & 5 & -2 \\ 6 & 3 & 9 & -4 \end{array}\right] X = \left[\begin{array}{c} 7 \\ 3 \\ 0 \\ 8 \end{array}\right]$$

$$\xrightarrow{R_1\left(\frac{1}{4}\right)} \begin{bmatrix} 1 & 3/4 & -7/4 & 1/4 \\ 2 & 5 & -7 & -3 \\ 8 & -3 & 5 & -2 \\ 6 & 3 & 9 & -4 \end{bmatrix} X = \begin{bmatrix} 7/4 \\ 3 \\ 0 \\ 8 \end{bmatrix}$$

$$\xrightarrow[R_4 - 6R_1]{R_2 - 2R_1,\ R_3 - 8R_1} \begin{bmatrix} 1 & 3/4 & -7/4 & 1/4 \\ 0 & 7/2 & -7/2 & -7/2 \\ 0 & -9 & 19 & -4 \\ 0 & -3/2 & 39/2 & -11/2 \end{bmatrix} X = \begin{bmatrix} 7/4 \\ -1/2 \\ -14 \\ -5/2 \end{bmatrix}$$

$$\xrightarrow{\left(\frac{2}{7}\right)R_2} \begin{bmatrix} 1 & 3/4 & -7/4 & 1/4 \\ 0 & 1 & -1 & -1 \\ 0 & -9 & 19 & -4 \\ 0 & -3/2 & 39/2 & -11/2 \end{bmatrix} X = \begin{bmatrix} 7/4 \\ -1/7 \\ -14 \\ -5/2 \end{bmatrix}$$

$$\xrightarrow[R_4 + \frac{3}{2}R_2]{R_1 - \frac{3}{4}R_2,\ R_3 + 9R_2} \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 10 & -13 \\ 0 & 0 & 8 & -7 \end{bmatrix} X = \begin{bmatrix} 13/7 \\ -1/7 \\ -107/7 \\ -19/7 \end{bmatrix}$$

$$\xrightarrow{R_3\left(\frac{1}{10}\right)} \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -13/10 \\ 0 & 0 & 18 & -7 \end{bmatrix} X = \begin{bmatrix} 13/7 \\ -1/7 \\ -107/70 \\ -19/7 \end{bmatrix}$$

$$\xrightarrow[R_4 - 18R_3]{R_1 + R_3,\ R_2 + R_3} \begin{bmatrix} 1 & 0 & 0 & -3/10 \\ 0 & 1 & 0 & -23/10 \\ 0 & 0 & 1 & -13/10 \\ 0 & 0 & 0 & 164/10 \end{bmatrix} X = \begin{bmatrix} 23/70 \\ -117/70 \\ -107/70 \\ 1736/70 \end{bmatrix}$$

$$\xrightarrow{\left(\frac{10}{184}\right)R_4 \rightarrow} \begin{bmatrix} 1 & 0 & 0 & -3/10 \\ 0 & 1 & 0 & -23/10 \\ 0 & 0 & 1 & -13/10 \\ 0 & 0 & 0 & 1 \end{bmatrix} X = \begin{bmatrix} 23/70 \\ -117/70 \\ -107/70 \\ 434/287 \end{bmatrix}$$

$$\xrightarrow[R_3 - 5R_1]{R_2 - 3R_1} A|B = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -8 & -1 & -1 \\ 0 & -8 & -1 & -1 \end{array}\right]$$

$$\xrightarrow{R_3 - R_2} A|B = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & -8 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\therefore \quad \rho(A) = \rho(A|B) = 2$

$\therefore$ By Gauss-Elimination method.

$$AX = B$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 3 & -5 & 2 \\ 5 & -3 & 4 \end{bmatrix} X = \begin{bmatrix} 3 \\ 8 \\ 14 \end{bmatrix}$$

$$\xrightarrow[R_3 - 5R_1]{R_2 - 3R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -8 & -1 \\ 0 & -8 & -1 \end{bmatrix} X = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$$

$$\xrightarrow{R_3 - R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -8 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$$

$$x + y + z = 3$$

$$-8y - z = -1$$

Take $z = t$

$$-8y - t = -1$$

$\therefore \quad -8y = -1 + t$

$\therefore \quad y = t - 1/8$

$$x + y + z = 3 \Rightarrow x + \frac{t-1}{8} + t = 3$$

$$x + \frac{t-1}{8} = 3 - t$$

$$x = 3 - t - \frac{t-1}{8} = \frac{3 - 8t - t - 1}{8} = 3 - \frac{3t - 1}{8}$$

$\therefore$ The solution is $\begin{pmatrix} (24 - 9t + 1)/8 \\ (t - 1)/8 \\ t \end{pmatrix}$

9. $2x - 3y + z = 0$

$x + 3y - 3z = 0$

$4x - y - 2z = 0$

Solution : The given system of equations in matrix form is as follows :

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 3 & -3 \\ 4 & -1 & -2 \end{bmatrix}$$

$$AX = 0$$

$\therefore$ By Gauss elimination method :

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 2 & -3 \\ 4 & -1 & -2 \end{bmatrix}$$

$$R_2 \leftrightarrow R_1 \rightarrow \begin{bmatrix} 1 & 2 & -3 \\ 2 & -3 & 1 \\ 4 & -2 & -2 \end{bmatrix} \xrightarrow[R_3 - 4R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & -3 \\ 0 & -7 & 7 \\ 0 & -9 & 10 \end{bmatrix}$$

$$\left(-\frac{1}{7}\right) R_2 \rightarrow \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & -1 \\ 0 & -9 & 10 \end{bmatrix} \xrightarrow{R_3 + 9R_2} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$\therefore$ $x + 2y - 3z = 0$

$y - z = 0$

$z = 0$

and $x + 2y - 3z = 0$

$x = 0$

---

10. $x + 2y + 3z = 0$

$2x + 3y + 4z = 0$

$7x + 13y + 9z = 0$

Solution : The given system of equations in matrix form is as follows :

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 7 & 13 & 9 \end{bmatrix}$$

$$\xrightarrow[R_3 - 7R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & -1 & -12 \end{bmatrix} \xrightarrow{R_3 - R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & -10 \end{bmatrix}$$

$$\therefore \quad x + 2y + 3z = 0$$
$$-y - 2z = 0$$
$$-10z = 0$$
$$z = 0$$

and $\quad -y - 2z = 0 \Rightarrow y = 0$

$$x + 2y + 3z = 0 \Rightarrow z = 0$$

---

11. $\quad x + y + 2z - 3w = 0$

$$3x + 2y - 4z + w = 0$$
$$5x - 3y + 2z + 6w = 0$$

The system of equations in matrix form is as follows :

$$A = \begin{bmatrix} 1 & -1 & 2 & -3 \\ 3 & 2 & -4 & 1 \\ 5 & -3 & 2 & 6 \end{bmatrix}$$

$$\xrightarrow[R_3 - 5R_1]{R_2 - 3R_1} \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 5 & -10 & 10 \\ 0 & 2 & -8 & 21 \end{bmatrix}$$

$$\left(\frac{1}{5}\right) R_2 \longrightarrow \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 2 & -8 & 21 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & -1 & 2 & -3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & -4 & 17 \end{bmatrix}$$

$$x - y + 2z - 3w = 0$$
$$y - 2z + 2w = 0$$
$$-4z + 17w = 0$$

Take $z = t$

$$\therefore \quad -4z + 17w = 4t + 17w = 0$$
$$17w = 4t$$
$$w = \frac{4t}{17}$$

and $\quad y - 2z + 2w = y - 2t + 2\frac{4t}{17} = 0$

$$\therefore \ \therefore \quad y - \frac{26t}{17} = 0$$
$$y = \frac{26t}{17}$$

and $\quad x - y + 2z - 3w = x - \frac{26t}{17} + 2t - 34\frac{t}{17} = 0$

$$x - \frac{4t}{17} = 0$$

$$x = \frac{4t}{17}$$

$\therefore$ The solution is, $\begin{pmatrix} 46t/17 \\ 26t/17 \\ t \\ 4t/17 \end{pmatrix}$ for any $t \in R$.

---

12. $x + 3y + 4z + 7w = 0$
$2x + 4y + 5z + 8w = 0$
$3x + y + 2z + 3w = 0$

Solution : The given system of equations in the matrix form is as follows :

$$A = \begin{bmatrix} 1 & 3 & 4 & 7 \\ 2 & 4 & 5 & 8 \\ 3 & 1 & 2 & 3 \end{bmatrix} \xrightarrow[R_3 - 3R_1]{R_2 - 2R_1} \begin{bmatrix} 1 & 3 & 4 & 7 \\ 0 & -2 & -3 & -6 \\ 0 & -8 & -10 & -18 \end{bmatrix}$$

$$R_3 - 4R_2 \quad \therefore A = \begin{bmatrix} 1 & 3 & 4 & 7 \\ 0 & -2 & -3 & -6 \\ 0 & 0 & 2 & 14 \end{bmatrix}$$

$x + 3y + 4z + 7w = 0$
$-2y - 3z - 6w = 0$
$2z + 14w = 0$

Take $w = s$

and $2z + 14w = 0 \Rightarrow 2z + 14s = 0$
$2z = -14s$
$z = -7s$

Let $-2y - 3z - 6w \Rightarrow -2y + 21s - 6s = 0$
$-2y + 15s = 0$
$\therefore y = \frac{15s}{2}$

and $x + 3y + 4z + 7w = x + \frac{3.15}{2}s - 28s + 7s = 0$

$\therefore \quad x + \frac{3s}{2} = 0$

$\therefore \quad x = \frac{-3s}{2}$

$\therefore$ The solution is, $\begin{pmatrix} -3s/2 \\ 15s/2 \\ -7s \\ s \end{pmatrix}$ where $s \in R$.

●●●

APPENDIX B

Find inverse and rational Form of given Matrix (with Matlab)

Q. 1 Find inverse and rational Form of given Matrix

(i) $A = \begin{bmatrix} 4 & 3 & 3 \\ -1 & 0 & 1 \\ -4 & -4 & -3 \end{bmatrix}$

Solution : Matlab command : To find inverse of given matrix are as follows :

```
>> A=[4 3 3;-1 0 1;-4 -4 -3]
A =
    4    3    3
   -1    0    1
   -4   -4   -3
```

Now we calculate inverse of given matrix

```
>> inv(A)
ans =
    0.5714   -0.4286    0.4286
   -1.0000         0   -1.0000
    0.5714    0.5714    0.4286
```

However to get a rational form,

```
>> format rat
>> inv(A)
ans =
     4/7        -3/7         3/7
    -1           0          -1
     4/7         4/7         3/7
```

(ii) $\begin{bmatrix} 2 & -2 & 0 & -1 \\ 0 & 2 & 1 & 2 \\ 1 & -2 & 3 & -2 \\ 0 & 1 & 2 & 2 \end{bmatrix}$

Solution : Matlab command :

```
>> B= [2 -2 0 -1;0 2 1 2;1 -2 3 -2;0 1 2 2]
B =
    2   -2    0   -1
    0    2    1    2
    1   -2    3   -2
    0    1    2    2
```

```
>> inv(B)
ans =
   0.4706   0.7647   0.0588  -0.4706
  -0.1176   1.0588   0.2353  -0.8824
  -0.1176   0.0588   0.2353   0.1176
   0.1765  -0.5882  -0.3529   0.8235

>> format rat
>> inv(B)
ans =

     8/17      13/17      1/17      -8/17
    -2/17      18/17      4/17     -15/17
    -2/17       1/17      4/17       2/17
     3/17     -10/17     -6/17      14/17
```

●●●

APPENDIX C : g-INVERSE AND MP-g INVERSE

Q. 1 Find g-inverse and MP g-inverse in each of the following

(i) $\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[1 2;1 2]
A =

   1   2
   1   2

>> rank(A)
ans =
   1
>> B=A(1,1)

B =
   1

>> rank(B)
ans =
   1
```

```
>> a1=[0]
a1 =
    0

>> a2=[0 0]
a2 =
    0     0
>> i=inv(B)
i =
    1

>> G=[i a1;a2]
G =
    1    0
    0    0

>> A*G*A
ans =
    1    2
    1    2
```

(ii) $\begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2 & 2 & 4 \\ 3 & 6 & 1 & 2 \\ 4 & 8 & 2 & 4 \end{bmatrix}$

Solution : Matlab command:

```
>> B=[1 2 3 6;1 2 2 4;3 6 1 2;4 8 2 4]
B =
    1    2    3    6
    1    2    2    4
    3    6    1    2
    4    8    2    4

>> rank(B)
ans =
    2

>> A=B(3:4,2:3)
A =
    6    1
    8    2
```

```
>> rank (A)
ans =
    2

>> i=inv(A)
i =
   0.5000   -0.2500
  -2.0000    1.5000

>> a1=[0 0]
a1 =
    0     0

>> a2=[0 0 0 0]
a2 =
    0     0     0     0

>> a3=[a1;i;a1]
a3 =
        0         0
   0.5000   -0.2500
  -2.0000    1.5000
        0         0
```

Q. 1 (iii) $A = \begin{bmatrix} 1 & 2 & 3 \\ -3 & 1 & 7 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[1 2 3;3 1 7]
A =
    1     2     3
    3     1     7

>> rank(A)
ans =
    2

>> B=A(1:2,2:3)
B =
    2     3
    1     7
```

```
>> rank(B)
ans =
     2

>> i=inv(B)
i =
    0.6364   -0.2727
   -0.0909    0.1818

>> a1=[0 0]
a1 =
     0     0

>> G=[a1;i]
G =
         0         0
    0.6364   -0.2727
   -0.0909    0.1818

>> A*G*A
ans =
     1     2     3
     3     1     7
```

Q. 1 (iv) $A = \begin{bmatrix} 7 & -2 \\ 3 & 2 \\ 1 & 4 \\ -3 & 3 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[7 -2;3 2;1 4;-3 3]
A =
     7    -2
     3     2
     1     4
    -3     3
```

```
>> rank(A)
ans =
    2

>> B=A(2:3,1:2)
B =
    3    2
    1    4

>> rank(B)

ans =

    2
>> i=inv(B)
i =
   0.4000   -0.2000
  -0.1000    0.3000

>> a1=[0 0]
a1 =
>> G=[a1;i;a1]
G =
  0   0.4000   -0.2000    0
  0   -0.1000    0.3000    0

>> A*G*A
Ans =
        7    2
        3    2
        1    4
        -3   3
```

Q. 1 (v) $A = \begin{bmatrix} -2 & 1 & 7 & 3 \\ 2 & 3 & 4 & 0 \\ 2 & 1 & 7 & 5 \\ 4 & 1 & 1 & 1 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[-2 1 7 3;2 3 4 0;2 1 7 3;4 1 1 1]
A =
   -2    1    7    3
    2    3    4    0
    2    1    7    3
    4    1    1    1

>> rank(A)
ans =
    4

>> i=inv(A)
i =
   -0.2500    0.0000    0.2500   -0.0000
    0.6000    0.2000   -0.8000    0.6000
   -0.3250    0.1000    0.4750   -0.4500
    0.7250   -0.3000   -0.6750    0.8500

>> Rank(A)=4 which is fullrow/colum rank, hence inv(A)=gin(A)
```

Q. 1 (vi) $D = \begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 2 & 3 & 4 \\ 3 & 7 & 5 & 1 \end{bmatrix}$

Solution : Matlab command:

```
>> D=[2 4 6 8;1 2 3 4;3 7 5 1]
D =
    2    4    6    8
    1    2    3    4
    3    7    5    1

>> rank(D)
ans =
    2

>> A=D(2:3,2:3)
A =
    2    3
    7    5
```

```
>> rank(A)
ans =
    2

>> i=inv(A)
i =
  -0.4545   0.2727
   0.6364  -0.1818

>> a1=[0 0]
a1 =
    0    0

>> a2=[0 0 0 0]
a2 =
    0    0    0    0

>> a3=[a1;i;a1]
a3 =
       0        0
  -0.4545   0.2727
   0.6364  -0.1818
       0        0

>> G=[a2'a3]
G =
0        0        0
0  -0.4545   0.2727
 0   0.6364  -0.1818
 0        0   0
>>D*G*D
Ans =  2.0000  5.0000  6.0000  8.0000
       1.0000  2.0000  3.0000  4.0000
       3.0000  7.0000  5.0000  1.0000
```

Find MP-g inverse each of the following :

Q.1 $A = \begin{bmatrix} 7 & -2 \\ 3 & 2 \\ 1 & 4 \\ -3 & 3 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[7 -2;3 2;1 4;-3 3]
A =
    7   -2
    3    2
    1    4
   -3    3

>> r=rank(A)
r =
    2

>> A1=A
A1 =
    7   -2
    3    2
    1    4
   -3    3

>> A2=[1 4;3 2;7 -2;-3 3]
A2 =
    1    4
    3    2
    7   -2
   -3    3

>> A2(2,:)=A2(2,:)-A2(1,:)*3
A2 =
    1    4
    0  -10
    7   -2
   -3    3

>> A2(3,:)=A2(3,:)-A2(1,:)*7
A2 =
    1    4
    0  -10
    0  -30
   -3    3
```

```
>> A2(4,:)=A2(4,:)+A2(1,:)*3
A2 =
    1     4
    0   -10
    0   -30
    0    15

>> A2(4,:)=A2(4,:).*2+A2(3,:)
A2 =
    1     4
    0   -10
    0   -30
    0     0

>> A2(3,:)=A2(3,:)-A2(2,:)*3
A2 =
    1     4
    0   -10
    0     0
    0     0

>> L=[1 4;0 -10]
L =
    1     4
    0   -10

>> I=[1 0 0 1;0 1 0 0;0 0 1 0;0 0 0 1]
I =
    1     0     0     1
    0     1     0     0
    0     0     1     0
    0     0     0     1

>> I(3,:)=I(3,:)+I(2,:)*3
I =
    1     0     0     1
    0     1     0     0
    0     3     1     0
    0     0     0     1
```

```
>> l(4,:)=l(4,:)*2-l(3,:)
l =
    1    0    0    1
    0    1    0    0
    0    3    1    0
    0   -3   -1    2

>> l(4,:)=l(4,:)-l(1,:)*3
l =
    1    0    0    1
    0    1    0    0
    0    3    1    0
   -3   -3   -1   -1

>> l(3,:)=l(3,:)+l(1,:)*7
l =
    1    0    0    1
    0    1    0    0
    7    3    1    7
   -3   -3   -1   -1

>> l(2,:)=l(2,:)+l(1,:)*3
l =
    1    0    0    1
    3    1    0    3
    7    3    1    7
   -3   -3   -1   -1
>> l1=[-7 -3 -1 -7;3 1 0 3;1 0 0 1;-3 -3 -1 -1]
L1 =
   -7   -3   -1   -7
    3    1    0    3
    1    0    0    1
   -3   -3   -1   -1

>> K=[-7 -3;3 1;1 0;-3 -3]
K =
   -7   -3
    3    1
    1    0
   -3   -3
```

```
>> K*L
ans =
   -7    2
    3    2
    1    4
   -3   18

ans =
   -7    2
    3    2
    1    4
   -3    3
```

Q.2 $A = \begin{bmatrix} -2 & 1 & 7 & 3 \\ 2 & 3 & 4 & 0 \\ 2 & 1 & 7 & 3 \\ 4 & 1 & 1 & 1 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[-2 1 7 3;2 3 4 0;2 1 7 3;4 1 1 1]
A =
   -2    1    7    3
    2    3    4    0
    2    1    7    3
    4    1    1    1

>> r=rank(A)
r =
    4

>> r=4
r =
    4

>> L=A
L =
   -2    1    7    3
    2    3    4    0
    2    1    7    3
    4    1    1    1
```

```
>> I=[1 0 0 0;0 1 0 0;0 0 1 0;0 0 0 1]
I =
    1    0    0    0
    0    1    0    0
    0    0    1    0
    0    0    0    1

>> K=I
K =
    1    0    0    0
    0    1    0    0
    0    0    1    0
    0    0    0    1

>> K+L
ans =
   -1    1    7    3
    2    4    4    0
    2    1    8    3
    4    1    1    2

>> T1=K'
T1 =
    1    0    0    0
    0    1    0    0
    0    0    1    0
    0    0    0    1

>> T2=L'
T2 =
   -2    2    2    4
    1    3    1    1
    7    4    7    1
    3    0    3    1

>> C=T1*A*T2
C =
   63   27   55    3
   27   29   35   15
   55   35   63   19
    3   15   19   19
```

```
>> d=inv(C)
d =
   1.0538   -0.1300   -1.1863    1.1225
  -0.1300    0.1400    0.0900   -0.1800
  -1.1863    0.0900    1.3838   -1.2675
   1.1225   -0.1800   -1.2675    1.2850

>> M=T2*d*T1
M =
  -0.2500   -0.0000    0.2500   -0.0000
   0.6000    0.2000   -0.8000    0.6000
  -0.3250    0.1000    0.4750   -0.4500
   0.7250   -0.3000   -0.6750    0.8500

>> A*M*A
ans =
  -2.0000    1.0000    7.0000    3.0000
   2.0000    3.0000    4.0000    0.0000
   2.0000    1.0000    7.0000    3.0000
   4.0000    1.0000    1.0000    1.0000
```

Q. 3 $A = \begin{bmatrix} 2 & 4 & 6 & 8 \\ 1 & 2 & 3 & 4 \\ 3 & 7 & 5 & 1 \end{bmatrix}$

Solution : Matlab command:

```
>> A=[2 4 6 8;1 2 3 4;3 7 5 1]
A =
    2    4    6    8
    1    2    3    4
    3    7    5    1

>> r=rank(A)
r =
    2

>> A1=A
A1 =
    2    4    6    8
    1    2    3    4
    3    7    5    1
```

```
>> A1(1,:)=A1(1,:)-A1(2,:)
A1 =
    1    2    3    4
    1    2    3    4
    3    7    5    1

>> A1(2,:)=A1(2,:)-A1(1,:)
A1 =
    1    2    3    4
    0    0    0    0
    3    7    5    1

>> L=[1 2 3 4;3 7 5 1]
L =
    1    2    3    4
    3    7    5    1

>> I=[1 0 0;0 0 1;0 0 1]
I =
    1    0    0
    0    0    1
    0    0    1

>> I1=[1 0 0;0 0 -1;0 1 0]
I1 =
    1    0    0
    0    0   -1
    0    1    0

>> I1(2,:)=I1(2,:)+I1(1,:)
I1 =
    1    0    0
    1    0   -1
    0    1    0

>> I1(1,:)=I1(1,:)+I1(2,:)
I1 =
    2    0   -1
    1    0   -1
    0    1    0
```

```
>> K=[2 0;1 0;0 1]
K =
    2    0
    1    0
    0    1

>> K*L
ans =
    2    4    6    8
    1    2    3    4
    3    7    5    1

>> T1=K'
T1 =
    2    1    0
    0    0    1

>> T2=L'
T2 =
    1    3
    2    7
    3    5
    4    1

>> C=T1*A*T2
C =
   150   180
    36    84

>> d=inv(C)
d =
    0.0137   -0.0294
   -0.0059    0.0245

>> M=T2*d*T1
M =
   -0.0078   -0.0039    0.0441
   -0.0275   -0.0137    0.1127
    0.0235    0.0118    0.0343
    0.0980    0.0490   -0.0931
```

```
>> A*M*A
ans =
   2.0000   4.0000   6.0000   8.0000
   1.0000   2.0000   3.0000   4.0000
   3.0000   7.0000   5.0000   1.0000
```

APPENDIX D

Eigen Value, Eigen Vectros, Spectral Decomposition and Power of a Matrix

Q.1 Find all the characteristic roots and characteristic vector corresponding to any non-zero characteristic root for the following matrix A.

(i) $A = \begin{bmatrix} 6 & 3 & -8 \\ 0 & -2 & 0 \\ 1 & 0 & -3 \end{bmatrix}$

Solution : Matlab Command :

```
>> A=[6 3 -8;0 -2 0;1 0 -3]
A =
     6     3    -8
     0    -2     0
     1     0    -3

>> [v,d]=eig(A)
v =
   257/259      985/1393     985/1393
     0          0            *
   257/2072     985/1393     985/1393

d =
     5     0     0
     0    -2     0
     0     0    -2

>> x1=v(:,1)
x1 =
   257/259
     0
   257/2072

>> x2=v(:,2)
x2 =
   985/1393
     0
   985/1393
```

Q.2 Express the following matrix A as a linear combination of idempotent matrx. Hence obtain A^4.

(i) $A = \begin{bmatrix} 3 & -1 & 4 \\ -1 & 0 & 1 \\ 4 & 1 & 2 \end{bmatrix}$

Solution : Matlab Command :

```
>> A=[3 -1 4;-1 0 1;4 1 2]
A =
     3    -1     4
    -1     0     1
     4     1     2

>> [v,d]=eig(A)
v =
   577/1008    1142/3475    1383/1841
  1922/3757     -372/433     -47/3362
   -289/451    -847/2159    1203/1823

d =
  -2215/934         0           0
      0        2447/2916        0
      0            0        3129/479

>> L1=d(1,1)
L1 =
  -2215/934

>> L2=d(2,2)
L2 =
   2447/2916

>> L3=d(3,3)
L3 =
   3129/479

>> x1=v(:,1)
x1 =
   577/1008
  1922/3757
   -289/451
```

```
>> x2=v(:,2)
x2 =
   1142/3475
   -372/433
   -847/2159

>> x3=v(:,3)
x3 =
   1383/1841
    -47/3362
   1203/1823

>> A=L1*x1*x1'+L2*x2*x2'+L3*x3*x3'
A =
     3          -1          4
    -1           *          1
     4           1          2

>> A4=(L1)^4*x1*x1'+(L2)^4*x2*x2'+(L3)^4*x3*x3'
A4 =
   1038         -10         891
    -10           9         -27
    891         -27         806

>> A^4
ans =
   1038         -10         891
    -10           9         -27
    891         -27         806
```

(ii) $B = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 4 & 2 \\ 1 & 5 & 3 \end{bmatrix}$

Solution : Matlab Command :

```
>> B=[1 2 -1;3 4 2;1 5 3]
B =
     1          2          -1
     3          4           2
     1          5           3
```

```
>> [v,d]=eig(B)
v =
   1114/15461     621/1108      434/643
    399/655       -85/5294     -821/1473
    849/1075     -987/1192     1216/2515

d =
   4926/709         0             0
      0          651/269          0
      0             0         -1082/791

>> [v1,d1]=eig(B')
v1 =
   1374/2185    -2203/2986     3097/6599
   -282/439     -839/1677      937/1129
    959/2189     489/1079      402/1333

d1 =
  -1082/791         0             0
      0          651/269          0
      0             0          4926/709

>> E1=v(:,1)/(v1(:,3)'*v(:,1))
E1 =
    331/3572
    539/688
    194/191

>> E2=v(:,3)/(v1(:,1)'*v(:,3))
E2 =
   1934/2849
   -833/1486
    124/255

>> E3=v(:,3)/(v1(:,1)'*v(:,3))
E3 =
   1934/2849
   -833/1486
    124/255
```

```
>> b1=E1*v1(:,3)'
b1 =
    175/4024      354/4603      236/8445
    389/1058      816/1255     3040/12867
   3077/6455     3833/4547      495/1616

>> b22=E2*v1(:,2)'
b22 =
   -907/1811    -2287/6734      511/1661
    902/2181     723/2578     -565/2224
   -649/1809     -507/2084      571/2591

>> b3=E3*v1(:,1)'
b3 =
    467/1094     -474/1087     2091/7031
   -472/1339      480/1333     -987/4019
     37/121     -442/1415      134/629

>> B=d(1,1)*b1+d(2,2)*b22+d(3,3)*b3
B =
  -1567/1049     1151/3726      659/1239
    537/133      1825/388      2961/2173
   1993/984      1589/279      1396/589

>> B4=(d(1,1)^4)*b1+(d(2,2)^4)*b22+(d(3,3)^4)*b3
B4 =
   6424/75      10792/65      6674/87
  14785/17      56461/37      31917/59
  16493/15      13684/7       70041/97

>> B^4
ans =
   6424/75      10792/65      6674/87
  14785/17      56461/37      31917/59
  16493/15      13684/7       70041/97
```

(iii) $a = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 4 & -4 & 5 \end{bmatrix}$

Solution : Matlab Command :

```
a=[1 2 -1;2 4 -2;4 -4 5]
a =
    1    2   -1
    2    4   -2
    4   -4    5

>> [v,d]=eig(a)
v =
   0.3714   -0.3333   -0.3333
  -0.5571   -0.6667   -0.6667
  -0.7428   -0.6667    0.6667

d =
  -0.0000         0         0
        0    3.0000         0
        0         0    7.0000

>> [v1,d1]=eig(a')
v1 =
  -0.1849    0.8944    0.8944
   0.7396   -0.4472    0.0000
  -0.6472   -0.0000    0.4472

d1 =
   7.0000         0         0
        0    0.0000         0
        0         0    3.0000

>> x1=v(:,1)
x1 =
   0.3714
  -0.5571
  -0.7428
```

```
>> x2=v(:,2)
x2 =
   -0.3333
   -0.6667
   -0.6667

>> x3=v(:,3)
x3 =
   -0.3333
   -0.6667
    0.6667

>> L1=d(1,1)
L1 =
-4.4409e-016

>> L2=d(2,2)
L2 =
    3.0000

>> L3=d(3,3)
L3 =
    7.0000

>> A=L1*x1*x1'+L2*x2*x2'+L3*x3*x3'
A =
    1.1111    2.2222   -0.8889
    2.2222    4.4444   -1.7778
   -0.8889   -1.7778    4.4444

>> A=(L1)^4*x1*x1'+(L2)^4*x2*x2'+(L3)^4*x3*x3'
A =
  1.0e+003 *
    0.2758    0.5516   -0.5156
    0.5516    1.1031   -1.0311
   -0.5156   -1.0311    1.1031
```

```
>> A^4
ans =
  1.0e+013 *
    0.3724    0.7449   -0.7391
    0.7449    1.4898   -1.4782
   -0.7391   -1.4782    1.4668
```

(iv) $B = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 3 \end{bmatrix}$

Solution : Matlab Command :

```
B=[1 2 -1;2 4 -2;-1 -2 3]
B =
     1     2    -1
     2     4    -2
    -1    -2     3

>> B=[1 2 -1;2 4 -2;-1 -2 3]
B =
     1     2    -1
     2     4    -2
    -1    -2     3

>> [v,d]=eig(B)
v =
    0.8944   -0.2433    0.3753
   -0.4472   -0.4865    0.7505
         0   -0.8391   -0.5439

d =
         0         0         0
         0    1.5505         0
         0         0    6.4495

>> [v1,d1]=eig(B')
v1 =
    0.8944   -0.2433    0.3753
   -0.4472   -0.4865    0.7505
         0   -0.8391   -0.5439
```

```
d1 =
       0         0         0
       0    1.5505         0
       0         0    6.4495

>> x1=v(:,1)
x1 =
   0.8944
  -0.4472
       0

>> x2=v(:,2)
x2 =
  -0.2433
  -0.4865
  -0.8391

>> x3=v(:,3)
x3 =
   0.3753
   0.7505
  -0.5439

>> L1=d(1,1)
L1 =
    0

>> L2=d(2,2)
L2 =
   1.5505

>> L3=d(3,3)
L3 =
   6.4495

>> B=L1*x1*x1'+L2*x2*x2'+L3*x3*x3'
B =
   1.0000    2.0000   -1.0000
   2.0000    4.0000   -2.0000
  -1.0000   -2.0000    3.0000
```

```
>> A=(L1)^4*x1*x1'+(L2)^4*x2*x2'+(L3)^4*x3*x3'
A =
  244.0000  488.0000 -352.0000
  488.0000  976.0000 -704.0000
 -352.0000 -704.0000  516.0000

>> B^4
ans =
  244.0000  488.0000 -352.0000
  488.0000  976.0000 -704.0000
 -352.0000 -704.0000  516.0000
```

Q. 3 $B = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 1 & 1 \\ -7 & 2 & -3 \end{bmatrix}$

Obtain all characteristic roots of the matrix and left characteristic vector corresponding to first characteristic roots. And right characteristic vector corresponding to a second and third characteristic roots.

Solution : Matlab Command :

```
>> B=[2 2 0;2 1 1;-7 2 -3]
B =
    2    2    0
    2    1    1
   -7    2   -3

>> [v1,d1]=eig(B)
v1 =
   0.0747   -0.6667   -0.4364
  -0.2242   -0.3333    0.2182
   0.9717    0.6667    0.8729

d1 =
  -4.0000         0         0
        0    3.0000         0
        0         0    1.0000

>> B'
ans =
    2    2   -7
    2    1    2
    0    1   -3
```

```
>> [v,d]=eig(B')
v =
  -0.6350    0.2357    0.7276
  -0.7620   -0.9428   -0.4851
  -0.1270   -0.2357    0.4851

d =
   3.0000         0         0
        0    1.0000         0
        0         0   -4.0000
For First  char root 4 left char vector is x1

>> x1=v1(:,1)
x1 =
   0.0747
  -0.2242
   0.9717
For Second char root 3 right char vector is x2
>> x2=v1(:,2)
x2 =
  -0.6667
  -0.3333
   0.6667
For Third char root 1 right char vector is x3

>> x3=v1(:,3)
x3 =
  -0.4364
   0.2182
   0.8729
```

APPENDIX E

Classification and Reduction of Quadratic Forms

Q.1 Examine the nature of quadratic Form

$230x^2 + 50xy + 94xz + 61y^2 - 8yz + 60z^2$

Obtain non-singular transformation which reduces it to canonical form.

Solution :

```
>> [230 25 47; 25 61 -4;47 -4 60]
ans =
  230   25   47
   25   61   -4
   47   -4   60
```

```
>> [v,d]=eig(a)
v =
   -0.3079    0.0313   -0.9509
    0.5224    0.8409   -0.1414
    0.7952   -0.5403   -0.2752

d =
   39.1769         0         0
         0   63.5011         0
         0         0  220.3220
```

Canonical Form :

```
>> D=inv(v)*a*v
D =
   39.1769         0    0.0000
    0.0000   63.5011    0.0000
    0.0000    0.0000  220.3220
```

Conclusion : Quadratic form is positive defininte.

Q.2 Examine the nature of quadratic Form

$2x^2 + y^2 + z^2 + 4xy - 7xz + 2yz$

Obtain non-singular transformation which reduces it to canonical form.

Solution :

```
>> a=[2 2 -7/2;2 1 1;-7/2 1 1]
a =
    2.0000    2.0000   -3.5000
    2.0000    1.0000    1.0000
   -3.5000    1.0000    1.0000

>> [v,d]=eig(a)
v =
    0.6181   -0.1232   -0.7764
   -0.4575   -0.8596   -0.2278
    0.6393   -0.4960    0.5877

d =
   -3.0999         0         0
         0    1.8637         0
         0         0    5.2362
```

Canonical Form :

```
>> D=inv(v)*a*v
D =
  -3.0999        0   -0.0000
  -0.0000    1.8637    0.0000
        0    0.0000    5.2362
```

Conclusion : Quadratic form is negative definite.

Q.3 (i) Find an orthogonal matrix of A such that B'AB is I.

$$B = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

(ii) Check whether matrix B is positive semi definite.

Solution :

```
>> a=[3 0 0;0 5 0;0 0 3]
a =
    3    0    0
    0    5    0
    0    0    3

>> d=eig(a)
d =
    3
    3
    5

>> v1=null(a-(3)*eye(3),'r')
v1 =
    1    0
    0    0
    0    1

>> v2=null(a-(5)*eye(3),'r')
v2 =
    0
    1
    0
```

>> v=[v1 v2]

Orthogonal Matrix

v =

1	0	0
0	0	1
0	1	0

To, check if v is a orthogonal matrix

>> I=v*v'

I =

1	0	0
0	1	0
0	0	1

Thus, V is a Orthogoanl Matrix

Canonical Form :

>> D=inv(v)*a*v

D =

3	0	0
0	3	0
0	0	5

Conclusion : As, all the eigen values are positive therefore, the quadratic form is positive definite.

❑❑❑

Model Question Paper – I

[Time : 3 Hours] **(For Non-Credit System)** **[Maximum Marks : 80]**

Instructions :

(i) Attempt any five questions.

(ii) Figures to the right indicate full marks.

(iii) Use of calculator and statistical tables is allowed.

(iv) Symbols and abbreviations have their usual meaning.

1. (a) **Prove or disprove :**

(i) For two matrices A, B; AB = 0 implies that either A = 0 or B = 0.

(ii) A square matrix A can be expresses as A = B + C where, B is a symmetric matrix and C is a skew-symmetric matrix.

(iii) If $A - A^T$ is a symmetric matrix then A = 0. **(6)**

(b) Define determinant of a matrix of order n. Further, if A and B are square matrices of the same order, then show that |AB| = |A| |B|. **(5)**

(c) If A and B are square matrices of order n, then show that **(5)**

$$\text{rank } (AB) \geq \text{rank } (A) + \text{rank } (B) - n.$$

2. (a) Explain the terms vector space and basis of vector space with the help of one example. **(5)**

(b) Show that the vectors $x_1 = (1, -1, 1)$ and $x_2 = (3, 4, -2)$ span the same vector space as that spanned by $y_1 = (2, 5, -3)$ and $y_2 = (9, 5, -1)$. Find an orthonormal basis for this vector space. **(6)**

(c) Let $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ and $\{\beta_1, \beta_2, \ldots, \beta s\}$ be two different bases of a vector space V. Show that s = k. **(4)**

3. (a) Find values of a and b for which following system of equations has (i) no solution, (ii) exactly one solution and (iii) infinitely many solutions.

$$\begin{aligned} -2x + bz &= 3 \\ ax + 2z &= 2 \\ 5x + 2y &= 1 \end{aligned}$$ **(6)**

(b) **Prove :**

(i) If A is an idempotent matrix, then A^n is also an idempotent matrix.

(ii) If A is an idempotent matrix, then rank (A) = trace A.

(iii) For two matrices, A of order $m \times n$ and B of order $n \times m$, trace (AB) = trace (BA). **(6)**

(c) Consider a system of linear equations AX = 0 where A is a matrix of order $m \times n$ with rank r. Show that the number of linearly independent solutions to the system are n – r. **(4)**

4. (a) Define (i) a generalised inverse and (ii) Moore-Penrose g-inverse.

Show that Moore-Penrose g-inverse is unique. **(6)**

(b) Obtain (i) a g-inverse and (ii) Moore-Penrose g-inverse of A, where A is given by

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix} \quad \textbf{(6)}$$

(c) Let $P = I - X(X'X)^{-} X'$ where $(X'X)^{-}$ is a g-inverse of X'X. Show that P is symmetric and idempotent. **(4)**

5. (a) (i) Define characteristic roots and characteristic vectors (left as well as right) of a square matrix.

(ii) Show that for a real symmetric matrix, the characteristic roots are all real. **(6)**

(b) Find characteristic roots of A where

$$a = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \quad 0 < \alpha,\ \beta < 1.$$

Also, find left as well as right characteristic vectors. **(6)**

(c) Let A be a matrix of order $m \times n$. Show that A has left-inverse L iff rank (A) = n. **(4)**

6. (a) State and prove Cayley-Hamilton theorem. **(4)**

(b) Show that characteristic vectors associated with distinct characteristic roots of a matrix are linearly independent. **(4)**

(c) Prove that characteristic roots of AB are same as characteristic roots of BA where, A is a $m \times n$ matrix and B is $n \times m$ matrix. **(4)**

(d) Find characteristic roots of the following matrix A.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 3 & 1 & 2 \\ -10 & 9 & 1 \end{bmatrix}. \quad \textbf{(4)}$$

7. (a) State and prove projection theorem. **(6)**

(b) Define inverse of a matrix. Is it unique? Find inverse of the following matrix A if it exists

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & 3 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad \textbf{(6)}$$

(c) **Prove or disprove :**

(i) $|A + B| = |A| + |B|$ if A and B are matrices of order n.

(ii) Let B an idempotent matrix of order n. B is either singular or identity matrix. **(4)**

8. (a) Define a quadratic form. Describe its usual classification.

Examine the nature of the following quadratic forms ;

(i) $Q_1 = x_1^2 + x_2^2 + 2x_3^2 + x_1x_2 + x_1x_3 - x_2x_3$

(ii) $Q_2 = x^2 + y^2 + z^2 - xy - yz - xz$. **(6)**

(b) State and prove a necessary and sufficient condition for a quadratic form to be positive definite. **(6)**

(c) If X'AX is a real quadratic form with rank (A) = r, show that there exists an orthogonal transformation X = PY such that X'AX is transformed to $\sum_j \lambda_i y_j^2$ where λ_j, j = 1, 2, …, r are the non-zero characteristic roots of A. **(4)**

•••

Model Question Paper – II

[Time : 3 Hours] **(For Non-Credit System)** **[Maximum Marks : 80]**

Instructions :

(i) Attempt any five questions.

(ii) Figures to the right indicate full marks.

(iii) Use of calculator and statistical tables is allowed.

(iv) Symbols and abbreviations have their usual meaning.

1. (a) Let A be a square matrix of order n. Show that it can be expressed as A = B + C where, B is a symmetric matrix of order n and C is a skew-symmetric matrix of order n. **(4)**

(b) Define an idempotent matrix. Prove that rank of an idempotent matrix is equal to its trace. **(4)**

(c) Define trace of a matrix of order n. Let A and B be two matrices of order n. Show that trace (AB) = trace (BA). **(4)**

(d) Prove or disprove :

(i) If A is a symmetric matrix then A^n is a symmetric matrix.

(ii) For two matrices A and B of appropriate orders AB = 0 implies that at least one of the matrices is a null matrix. **(2 + 2)**

2. (a) (i) Define the determinant of a square matrix.

(ii) Let A and B be two square matrices of order n. Show that |AB| = |A| |B|. **(2 + 4)**

(b) Define row rank and column rank of a matrix. Show that they are equal. **(2 + 4)**

(c) Prove that the value of the determinant of a matrix is unchanged if the multiple of one row (column) is added to another row (column) of the matrix. **(4)**

3. (a) Define an orthogonal matrix. Show that columns of an orthogonal matrix are linearly independent. However, the converse is not true. **(4)**

(b) Prove or disprove :

(i) Let A and B be idempotent matrices of order n, then A + B is idempotent.

(ii) If A is an orthogonal matrix then −A is orthogonal. **(2 + 2)**

(c) Prove that A is an idempotent matrix of order n iff r(A) + r(I − A) = n. **(4)**

(d) Find rank of the matrix A given below :

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 2 \\ 1 & -1 & 3 & 2 \end{bmatrix}. \quad \textbf{(4)}$$

4. (a) Illustrate the following terms by giving one example each :

(i) Linearly independent vectors.

(ii) Vector space.

(iii) Basis of a vector space.

(iv) Vector space spanned by vectors $\alpha_1, \alpha_2, \ldots, \alpha_k$. **(4)**

(b) Let $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ and $\{\beta_1, \beta_2, \ldots, \beta_s\}$ be two different bases of a linear vector space V. Show that k = s. **(4)**

(c) (i) Show that if $\alpha_1, \alpha_2, \ldots, a_k$ are orthogonal vectors, then they are linearly independent.

(ii) (A) Describe Gram-Schmidt orthogonalization procedure.

(B) Use this procedure to find an orthonormal basis for the vector space spanned by the following vectors.

$$\alpha_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \alpha_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \alpha_3 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \quad \textbf{(2 + 2 + 4)}$$

5. (a) Give two definitions of a g-inverse of a matrix and establish their equivalence. Using a suitable example, demonstrate that a g-inverse exists for a square matrix that is singular. **(3 + 3)**

(b) Define Moore-Penrose g-inverse of a matrix. Show that it is unique. **(4)**

(c) Define solution space and coefficient space of a homogeneous system of m linear equations in n unknowns (AX = 0).

Show that rank (solution space) = n – rank (coefficient space). **(6)**

6. (a) With the help of one example, explain the following terms :

(i) Characteristic root.

(ii) Right characteristic vector.

(iii) Left characteristic vector. **(4)**

(b) Obtain characteristic roots and corresponding right and left characteristic vectors for the following matrix A.

$$A = \begin{bmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}.$$

Using these obtain spectral decomposition of this matrix. Hence, find A^3.

(3 + 4 + 1)

(c) Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be characteristic roots of a square matrix A of order n. Show that (i) $\sum_{i=1}^{n} \lambda_i = \text{trace } (A)$, (ii) $\prod_{i=1}^{n} \lambda_i = |A|$. **(4)**

7. (a) Show that if A is a real symmetric matrix then there exists a real orthogonal matrix P such that P'AP = diag $(\lambda_1, \ldots, \lambda_n)$ where $\lambda_1, \ldots, \lambda_n$ are characteristic roots of A. **(6)**

(b) Investigate for which values of a and b, the following system of simultaneous equations has (i) no solution, (ii) unique solution and (iii) infinitely many solutions.

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + 3z &= 10 \\ x + 2y + az &= b \end{aligned}$$

(6)

(c) **Prove or disprove :**

(i) The set of all solutions of homogeneous system of m linear equations in n unknowns, $A\underline{X} = \underline{0}$ forms a linear vector space.

(ii) The non-homogeneous system of m linear equations in n unknowns, $A\underline{X} = \underline{b}$ is consistent iff $\underline{b}$ belongs to linear vector space spanned by columns of A.

(2 + 2)

8. (a) Define :

(i) A positive definite (p.d.) quadratic form.

(ii) A positive semidefinite (p.s.d.) quadratic form.

Also show that a quadratic form can always be represented by a symmetric matrix. **(4)**

(b) Prove or disprove :

For a symmetric matrix A,

(i) The quadratic form X'AX is p.s.d. if A is idempotent.

(ii) The qaudratic form X'AX is p.d.f if A is orthogonal.

(iii) The quadratic form $X'A^2X$ is always p.s.d. **(6)**

(c) Examine the nature of each of the following forms :

(i) $2x_1^2 + x_2^2 - 3x_3^2 - 8x_2x_3 - 4x_1x_3 + 12x_1x_2$

(ii) $6x_1^2 + 3x_2^2 + 3x_3^2 - 4x_1x_2 - 2x_2x_3 + 4x_1x_3$

(iii) $x_1^2 + 2x_2^2 + 3x_3^2 + 2x_2x_3 - 2x_3x_1 + 2x_1x_2$. **(2 + 2 + 2)**

•••

Model Question Paper – III

[Time : 3 Hours] **(For Non-Credit System)** **[Maximum Marks : 80]**

Instructions :

(i) Attempt any five questions.

(ii) Figures to the right indicate full marks.

(iii) Use of calculator and statistical tables is allowed.

(iv) Symbols and abbreviations have their usual meaning.

1. (a) Define : (i) Symmetric matrix and (ii) Skew-symmetric matrix.

Show that a square matrix can always be expressed as a sum of a symmetric and a skew-symmetric matrix. **(4)**

(b) Define an indempotent matrix and show that rank of an imdempotent matrix is equal to its trace. **(4)**

(c) Show that A'A = 0 implies that A = 0 where A is a square matrix of order n. Further, does $A^2 = 0$ imply that A = 0 ? Justify your answer. **(4)**

(d) Define trace of a matrix. Let A, B, C be 3 square matrices. Prove that trace (ABC) = trace (CBA). **(4)**

2. (a) Explain the terms (i) vector space and (ii) basis of vector space with one illustration each. Show that if $\{\underline{\alpha}_1, \underline{\alpha}_2, \ldots, \underline{\alpha}_k\}$ is a basis of vector space V, then $\{\underline{\alpha}_1, \underline{\alpha}_1 + \underline{\alpha}_2, \ldots \underline{\alpha}_1 + \underline{\alpha}_2 + \ldots + \underline{\alpha}_k\}$ is also a basis of V.

(b) Describe Gram-Schmidt orthogonalisation process. Using this construct an orthonormal basis for the vector space spanned by $\underline{\alpha}_1, \underline{\alpha}_2, \underline{\alpha}_3$ as given below.

$$\underline{\alpha}_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \quad \underline{\alpha}_2 = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix} \quad \underline{\alpha}_3 = \begin{pmatrix} 6 \\ 1 \\ 0 \end{pmatrix} \qquad \textbf{(8)}$$

3. (a) Define determinant of a matrix of order n, $n \geq 2$. Obtain the determinant of the following matrix.

$$A = \begin{bmatrix} 1 & \rho & \ldots\ldots & \rho \\ \rho & 1 & \ldots\ldots & \rho \\ : & & & \\ : & & & \\ \rho & \rho & \ldots\ldots & 1 \end{bmatrix}_{n \times n} \qquad \textbf{(5)}$$

(b) Prove :

(i) Any property of |A| in terms of rows of A is also true in terms of columns of A.

(ii) The value of determinant of a matrix is unchanged if the multiple of one column is added to another column of the matrix. **(6)**

(c) Define row-rank and column rank of a matrix and show that they are equal. **(5)**

4. (a) Prove : A is an idempotent matrix of order n

iff rank (A) + rank (I − A) = n. **(4)**

(b) Find rank of the following matrix.

$$A = \begin{bmatrix} 1 & -2 & -3 & -2 & 1 \\ 0 & 2 & 2 & 1 & -5 \\ 3 & -2 & 0 & -1 & -7 \\ 0 & 1 & 2 & 1 & -6 \end{bmatrix}. \quad \textbf{(4)}$$

(c) Let A and B be two matrices of order n.

Show that rank (AB) ≤ min (rank A), rank (B)). **(4)**

(d) Solve the following system of linear equations and obtain all possible solutions.

$$5x_1 + 3x_2 + 7x_3 = 4$$
$$3x_1 + 26x_2 + 2x_3 = 9$$
$$7x_1 + 2x_2 + 10x_3 = 5 \quad \textbf{(4)}$$

5. (a) Let AX = 0 be a system of m linear equations in n unknowns.

Prove that rank of solution space = n – rank of coefficient space. **(6)**

(b) Prove :

(i) A system of non-homogeneous equations AX = b has a solution iff rank of linear vector space spanned by $\alpha_1, \ldots, \alpha_n$, is same as rank of linear vector space spanned by $\alpha_1, \alpha_2, \ldots, \alpha_n$, b. where $\alpha_1, \alpha_2, \ldots \alpha_n$ are columns of A.

(ii) The set of all solutions of a system of non-homogeneous equations AX = b does not constitute a linear vector space. **(3 + 1)**

(c) Define g-inverse. Obtain two distinct g-inverse for the following matrix :

$$A = \begin{bmatrix} 5 & 2 & -1 \\ 3 & 2 & 3 \\ 3 & 2 & 3 \end{bmatrix} \quad \textbf{(6)}$$

6. (a) Define Moore-Penrose (MP) g-inverse. Show that it is unique. Further, obtain the same for the following matrix.

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}. \quad \textbf{(8)}$$

(b) (i) Define characteristic roots and characteristic vectors of a square matrix of order n.

(ii) Prove that characteristic roots of every real symmetric matrix are real. **(2 + 2)**

(c) State and prove Cayley-Hamilton theorem. **(4)**

7. (a) Define spectral decomposition of a matrix.

Using this, obtain A^4 if A is given by

$$A = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$ **(6)**

(b) (i) Define right inverse of a matrix A of order $m \times n$.

(ii) Show that $A_{m \times n}$ has right inverse iff rank (A) = m. **(6)**

(c) Define a quadratic form. When do you say that a quadratic form $Q(\underline{x})$ is non-negative definite (n.n.d.) ? Show that if A is symmetric, idempotent matrix then it is n.n.d. **(4)**

8. (a) Prove : If A is a real symmetric matrix, then there exists a real orthogonal matrix P such that P'AP = diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are characteristic roots of A. **(6)**

(b) State and prove necessary and sufficient condition for a quadratic form to be positive definite. **(6)**

(c) Identify the nature of following quadratic forms.

(i) $x_1^2 + 2x_2^2 + 3x_3^2 + 2x_2x_3 - 2x_1x_3 + 2x_1x_2$

(ii) $5x_1^2 + 26x_2^2 + 10x_3^2 + 4x_2x_3 + 14x_3x_1 + 6x_1x_2$. **(4)**

•••

Model Question Paper – IV

[Time : 3 Hours] **(For Credit System)** **[Maximum Marks : 80]**

Instructions :

(i) Question 1 is compulsory. Attempt any five equations from Question 2 to 9.

(ii) Figures to the right indicate full marks.

(iii) Use of calculator and statistical tables is allowed.

(iv) Symbols and abbreviations have their usual meaning.

1. (a) Choose the correct alternative for each of the following :

(1) Consider the system of equations $A\underline{x} = \underline{0}$, where A is $n \times n$ matrix and $\underline{X} = (x_1, x_2, \ldots, x_n)'$. If rank (A) < n then system has

(i) a non-trivial solution

(ii) unique solution

(iii) no solution

(iv) no non-trivial solution.

(2) Let A be $m \times n$ matrix over $\mathbb{R}$, which of the following statements is false ?

(i) Row rank of A < column rank of A

(ii) Row rank of $A \leq m$

(iii) Row rank of $A \leq n$

(iv) Row rank of A = column rank of A.

(3) The quadratic form corresponding to matrix $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ is

(i) positive definite

(ii) positive semidefinite

(iii) negative definite

(iv) indefinite.

(4) Let A be 3×3 matrix having eigen values 3, 2, −5 then tr (A) and det (A) are

(i) 3, −30 (ii) 0, − 30 (iii) −30, 0 (iv) −30, 3.

(5) The system of equations

$$ax - 10y = 8$$
$$10x - by = 2$$

has a unique solution if

(i) $ab \neq 100$ (ii) $ab = 100$ but $b \neq 2$ (iii) $ab = 100$ but $b \neq 2$ and $a \neq 50$

(iv) $ab = 100$ but $a \neq 50$.

(b) State whether each of the following statements is true or false.

(i) Basis of a vector space is unique.

(ii) The two matrices A and $P^{-1}AP$ have the same characteristic roots, where P is non-singular matrix.

(c) Explain the following terms :

(i) Linearly dependent set of vectors.

(ii) Elementary matrix.

(iii) Algebraic multiplicity of characteristic root. **(5 + 2 + 3)**

2. (a) Define rank of a matrix. Prove that rank of product of two matrices cannot exceed the rank of either matrix.

(b) Define determinant of an $n \times n$ matrix, $n \geq 2$. Obtain |A|, where

$$A = \begin{bmatrix} 1 & r & r^2 & r^3 \\ r & 1 & r & r^2 \\ r^2 & r & 1 & r \\ r^3 & r^2 & r & 1 \end{bmatrix}.$$

(c) Prove or disprove :

If A is idempotent matrix then A^n is also idempotent matrix. **(5 + 6 + 3)**

3. (a) Define trace of a matrix of order n. Let A and B be two matrices of order n. Show that

$$\text{trace (AB)} = \text{trace (BA)}$$

(b) Illustrate the following terms with one example of each.

(i) Vector space.

(ii) Basis of a vector space.

(iii) Dimension of a vector space.

(c) Define orthogonal vectors. Prove that $\alpha_1, \alpha_2, \ldots, \alpha_k$ are k non-null orthogonal vectors then they are linearly independent. **(4 + 6 + 4)**

4. (a) What is Gram-Schmidt orthogonalization ? Carry out this procedure on the following vectors and get an orthogonal basis for the vector space spanned by these vectors.

$$\underline{\alpha}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \underline{\alpha}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \underline{\lambda}_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

(b) Consider the system of equations $A\underline{x} = \underline{0}$ where, A is $m \times n$ matrix with rank (A) = r. Show that there are $n - r$ linearly independent solutions to this system. **(8 + 6)**

5. (a) Determine conditions under which the non-homogeneous system $A\underline{x} = \underline{b}$ of m equations in n-unknowns has (i) no solution (ii) a unique solution (iii) infinite number of solutions.

(b) Solve the following system of linear equations and obtain all possible solutions.

$$-x_1 + 3x_2 + 3x_3 + 2x_4 = 0$$

$$2x_1 + 6x_3 + x_3 = 0$$

$$-2x_1 + 4x_2 + 2x_3 + 4x_4 = 0$$

(c) Define left inverse of an $m \times n$ matrix. Further prove that an $m \times n$ matrix right inverse R iff rank (A) = m. **(4 + 4 + 6)**

6. (a) Define :

(i) A generalized inverse.

(ii) Moore-Penrose g-inverse.

Show that Moore-Penrose g-inverse is unique.

(b) Obtain :

(i) Two distinct g-inverse of A.

(ii) Moore-Penrose g-inverse of A.

where, A is given by

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix}. \qquad \textbf{(6 + 8)}$$

7. (a) (i) Define characteristic roots and characteristic vectors (left as well as right) of a square matrix.

(ii) Prove that characteristic roots of every real symmetric matrix are real.

(b) State and prove Cayley-Hamilton theorem.

(c) Prove that

(i) If A is non-singular matrix then every characteristic root of matrix A is non-zero.

(ii) If λ is characteristic root of non-singular matrix A then $1/\lambda$ is characteristic root of A^{-1}. **(6 + 4 + 4)**

8. (a) Prove that if A is real symmetric matrix, then there exists a real orthogonal matrix P such that P'AP = diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are characteristic roots of A.

(b) Define a quadratic form. Describe it's usual classification. Identify the nature of the following quadratic forms.

(i) $x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_2x_3 + 6x_1x_3$

(ii) $2x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 - 2x_1x_3 - 4x_1x_3$

(iii) $x_1^2 + 4x_1x_2 + 4x_2^2 - 4x_3^2$. **(6 + 8)**

9. (a) State and prove the necessary and sufficient condition for a quadratic form to be positive definite.

(b) Prove that two characteristic vectors corresponding to two distinct roots of a real symmetric matrix are orthogonal.

(c) Prove or disprove :

(i) The quadratic form X'AX is positive semi-definite if A is idempotent.

(ii) The quadratic form X'AX is positive definite if A is orthogonal. **(6 + 4 + 4)**

●●●

Model Question Paper – V

[Time : 3 Hours] **(For Credit System)** **[Maximum Marks : 80]**

Instructions :

(i) Question 1 is compulsory. Attempt any five equations from Question 2 to 9.

(ii) Figures to the right indicate full marks.

(iii) Use of calculator and statistical tables is allowed.

(iv) Symbols and abbreviations have their usual meaning.

1. (a) Choose the correct alternative for each of the following :

(1) Let $\{e_1, e_2, \ldots, e_n\}$ be a standard basis of $\mathbb{R}^n$ then $x_1 = a_{11}e_1$, $x_2 = a_{12}e_1 + a_{22}e_2$, …, $x_n = a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n$ is a basis of $\mathbb{R}^n$ iff

(i) $a_{11} \cdot a_{22} \ldots a_{nn} = 0$, (ii) $a_{11} \cdot a_{22} \ldots a_{nn} \neq 0$, (iii) $a_{nn} = 1$, (iv) $a_{11} = 1$.

(2) Consider a matrix $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{bmatrix}$ then unity of matrix A is

(i) 4, (ii) 3, (iii) 2, (iv) 1.

(3) If matrix A has characteristic polynomial f(x) then transpose of a matrix A has characteristic polynomial.

(i) $-f(x)$, (ii) $f(x)$, (iii) $f(-x)$, (iv) $\frac{1}{f(x)}$.

(4) Let V be real vectors sace and $\{x_1, x_2, x_3\}$ be a basis for V, then which of the following statements are true.

(a) $\{x_1, x_2, x_3\}$ is a basis for V.

(b) dimension of V is 3.

(c) $\{x_1, x_2, x_3\}$ is a orthogonal basis.

(d) $\{X_1 - X_2, X_2 - X_3, X_1 - X_3\}$ is a basis for V.

(i) Statement (a), (b), (c) are correct.

(ii) Statements (a), (b), (d) are correct.

(ii) Statement (a), (b) are correct.

(iv) All statements are correct.

(5) The quadratic form corresponding to a matrix $\begin{bmatrix} 1 & 4 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

(i) positive definite, (ii) positive semidefinite, (iii) negative definite, (iv) indefinite. **(5)**

(b) Illustrate the following terms with one example each :

(i) Elementary matrix

(ii) Linearly dependent set of vectors

(iii) Orthonormal vectors. **(3)**

(c) Prove or disprove :

(i) If A is symmetric matrix then A^n is also symmetric matrix.

(ii) If A is skew-symmetric non-singular matrix then A^{-1} is also skew symmetric matrix.

(iii) If A is idempotent matrix then A^n is also idempotent matrix.

(iv) If A is orthogonal matrix then $-A$ is orthogonal matrix. **(8)**

2. (a) Show that $A'A = 0$ implies that $A = 0$, where A is a square matrix of order n. Further, does $A^2 = 0$ imply $A = 0$? Justify your answer. **(4)**

(b) (i) Find |A|, where A is

$$A = \begin{bmatrix} n^2 & (n+1)^2 & (n+2)^2 \\ (n+1)^2 & (n+2)^2 & (n+3)^2 \\ (n+2)^2 & (n+3)^2 & (n+4)^2 \end{bmatrix}$$

(ii) Describe the elementary transformations on a matrix and explain its effect on the rank of a matrix. **(4 + 4)**

(c) Define the rank of a matrix. Prove that the rank of a product of two matrices cannot exceed the rank of either matrix. **(4)**

3. (a) Define an orthogonal matrix. Show that the columns of an orthogonal matrix are linearly independent and the converse of this is not true. **(4)**

(b) Define row rank and column rank of a matrix and show that they are equal. **(4)**

(c) (i) Explain the terms vectors space and basis of a vector space with the help of an example.

(ii) Let $V = \{(x_1, x_2, 0, 0, \ldots, 0)_{1 \times n} \mid x_1, x_2 \neq 0, x_1, x_2 \text{ are real}\}$.

Let $e_1 = (1, 0, 0, \ldots, 0)$ and $e_2 = (0, 1, \ldots, 0)$. Examine whether V is linear vector space and $\{e_1, e_2\}$ form a basis of V. **(4 + 4)**

4. (a) Define orthogonal vectors and linearly independent vectors. Are orthogonal vectors linearly independent ? Justify your answer. **(4)**

(b) Prove or disprove :

(i) Subset of a linearly dependent set of vectors is linearly dependent.

(ii) If η is a linear combination of the set $\{x_1, x_2, \ldots, x_n\}$ then the set η, $x_1, x_2, x_2, \ldots, x_n\}$ is linearly independent. **(2 + 2)**

(c) (i) Determine the conditions under which non-homogeneous system $A\underline{X} = \underline{b}$ of m equations in n-unknowns has

(1) no solution, (2) a unique solution, (iii) infinite number of solutions.

(ii) Prove that a system of non-homogeneous linear equations $A\underline{X} = \underline{b}$ has a solution iff rank of linear vector space spanned by vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is same as that rank of linear vector space spanned by vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_n, \underline{b}\}$ where, $\alpha_1, \alpha_2, \ldots, \alpha_n$ are columns of A.

(iii) The set of all solutions of a system of homogeneous equations. $AX = 0$ constitute a linear vector space. **(3 + 3 + 2)**

5. (a) Use Gram-Schmidt orthogonalization procedure to find orthogonal basis for the vector space spanned by the following vectors

$$\alpha_1 = \begin{pmatrix}1\\2\\1\end{pmatrix}, \quad \alpha_2 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad \alpha_3 = \begin{pmatrix}3\\1\\0\end{pmatrix} \qquad \textbf{(4)}$$

(b) State and prove "Full rank factorization theorem". **(4)**

(c) If A is idempotent matrix of order $n \times n$ then show that :

(i) rank (A) = trace (A)

(ii) rank (A) + rank (I − A) = n.

Is the converse of (ii) true ? **(8)**

6. (a) Define a generalized inverse and the Moore-Penrose generalized inverse of a matrix. Obtain two generalized inverse for A where A = [1 2 3] and comment on the result. **(6)**

(b) Suppose $A = \underline{u}\underline{u}'$, $\underline{u} \neq 0$, $v \neq 0$. Obtain the Moore-Penrose generalized inverse of A in terms of $\underline{u}$ and $\underline{v}$. Hence or otherwise, obtain the Moore-Penrose generalized inverse of $\begin{bmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$. Comment on your finding. **(6)**

(c) **Prove or disprove :**

(i) Generalized inverse of an orthogonal matrix is unique.

(ii) If A is symmetric idempotent matrix, then A itself is its Moore-Penrose generalized inverse. **(2 + 2)**

7. (a) Obtain characteristic roots and corresponding right and left characteristic vector for the following matrix A

$$A = \begin{bmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}$$

Using these obtain spectral decomposition of this matrix. Hence, find A^3.

(3 + 4 + 1)

(b) Let A be n × n matrix with n-distinct characteristic roots $\lambda_1, \lambda_2, \ldots, \lambda_n$. Show that there exist a non-singular matrix S such that $A = S \cap S^{-1}$ where, $\cap = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Prove all necessary results. **(4)**

(c) Prove that :

(i) If λ is characteristic root of a non-singular matrix A, then $\frac{|A|}{\lambda}$ is a characteristic root of Ads (A).

(ii) The two matrices A, $P^{-1}AP$ have same characteristic roots, where P is non-singular matrix. **(4)**

8. (a) Define a quadratic form. Describe its usual classification. Examine the nature of the following quadratic forms :

(i) $Q_1 = x_1^2 + x_2^2 + 2x_3^2 + d_1x_2 + x_1x_3 - x_2x_3$

(ii) $Q_2 = x^2 + y^2 + z^2 - xy - yz - xz$. **(6)**

(b) Prove that two characteristic vectors corresponding to two distinct root of real symmetric matrix are orthogonal. **(4)**

(c) Prove or disprove :

For a symmetric matrix A,

(i) The quadratic form X'AX is positive semidefinite if A is an idempotent matrix.

(ii) The quadratic form X'AX is positive definite if A is an orthogonal matrix.

(iii) The quadratic form $X'A^2X$ is always positive semidefinite. **(6)**

•••

Model Question Paper – VI

[Time : 3 Hours] **(For Non-Credit System)** **[Maximum Marks : 80]**

Instructions :

(i) Attempt any five questions.

(ii) Figures to the right indicate full marks.

(iii) Use of calculator and statistical tables is allowed.

(iv) Symbols and abbreviations have their usual meaning.

1. (A) Explain the following terms with one example each :

(i) Elementary matrix. (ii) Skew-symmetric matrix.

(iii) Determinant of matrix. (iv) Inverse of matrix. **(8)**

(B) Define rank of a matrix. Prove that row rank and column rank of matrix are equal. **(4)**

(C) Define idempotent matrix. Show that for an idempotent matrix A, rank (A) = trace (A). **(4)**

2. (A) Given A is a square matrix of order n. Show that A'A = 0 implies that A = 0. Further, does $A^2 = 0$ imply A = 0 ? Justify your answer. **(4)**

(B) Define an orthogonal matrix. Show that the columns of an orthogonal matrix is linearly independent and converse of this is not true.

(C) Prove or disprove :

(i) If A is a symmetric matrix A^n is also symmetric.

(ii) If A and B are two idempotent matrices of order n, then (A − B) is idempotent matrix. **(4)**

(D) Define trace of a matrix of order n. Let $A_{m \times n}$ and $B_{n \times m}$ be two matrices.

Show that trace (AB) = trace (BA). **(4)**

3. (A) Define vector space and its basis with an example and prove that the number of members in any one basis of a subspace is the same in any other basis. **(8)**

(B) Suppose V_1 and V_2 are n-dimensional vector spaces. Check if $V_1 + V_2$, $V_1 \cup V_2$, $V_1 \cap V_2$ and V_1/V_2 are vector spaces. **(8)**

4. (A) Use Gram-Schmidt orthogonalisation procedure on the following vectors and get an orthogonal basis for the vector spanned y these vectors.

$\alpha_1 = [1, 1, 1, 1]$ $\alpha_2 = [1, 2, 3, 4]$ $\alpha_3 = [1, 2, 2, 1]$ $\alpha_4 = [1, 1, -2, 3]$ **(8)**

(B) Define linear independence of a set of vectors. Give one example of a set of independent vectors and one of dependent vectors with justification. **(4)**

(C) Prove or disprove :

(i) Subset of linearly dependent set of vectors is linearly dependent.

(ii) If n is linear combination of the set $\{x_1, x_2, x_3, \ldots\ldots, x_n\}$ then the set $\{n, x_1, x_2, x_n, \ldots, x_n\}$ is linearly independent.

5. (A) Solve the following system of linear equation and obtain all possible solutions, if it is consistent

$$X_1 + 2X_2 - 3X_3 - 4X_4 = 0$$
$$X_1 + 3X_2 - X_3 - 2X_4 = 0$$
$$2X_1 + 5X_2 - 2X_3 - 4X_4 = 0$$ **(6)**

(B) Define a generalized inverse and Moore-Penrose generalized inverse with an example. Show that M-P generalized inverse is unique. **(6)**

(C) Prove or disprove : Let A^- generalized inverse of a matrix A then

(i) AA^- is idempotent matrix.

(ii) Rank (A) ≤ Rank (A^-). **(4)**

6. (A) State and prove that Cayley Hamilton theorem. **(4)**

(B) Prove or disprove :

(i) If A is symmetric matrix then its left eigen vector is same as its right eigen vector.

(ii) If t is eigen value of matrix A then (t + A) is eigen value of (A + ct).

(iii) If A is orthogonal matrix then all eigen values of matrix A are + 1 or – 1. **(6)**

(C) Prove that if A is a real symmetric matrix of order n then there exists a real orthogonal matrix P such that P'AP is diagonal matrix with diagonal entries are $t_1, t_2, t_3, \ldots, t_n$ where, $t_1, t_2, t_3, \ldots, t_n$ are eigen values of A. **(6)**

7. (A) Define algebraic multiplicity and geometric multiplicity of an eigen value. Given an example with justification of a matrix having an eigen value with algebraic multiplicity 2 and geometric multiplicity 1. **(4)**

(B) Prove that for a real symmetric matrix eigen values are real. **(4)**

(C) Using spectral decomposition of a matrix, obtain A^8 if A is given by

$$A = \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{pmatrix}$$ **(8)**

8. (A) Define a quadratic form and describe its usual classification. Examine the nature of the following quadratic form.

$$xy + yz + xz$$ **(6)**

(B) Describe non-singular transformation in quadratic form. Show that a nature of a quadratic form is invariant under non-singular transformation. **(4)**

(C) State and prove that a necessary and sufficient condition for a quadratic form to be non-negative definite. **(6)**

●●●